



Universiteit  
Leiden  
The Netherlands

## **Analyzability and semantic associations in referring expressions : a study in comparative lexicology**

Urban, M.

### **Citation**

Urban, M. (2012, October 10). *Analyzability and semantic associations in referring expressions : a study in comparative lexicology*. Retrieved from <https://hdl.handle.net/1887/19940>

Version: Not Applicable (or Unknown)

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/19940>

**Note:** To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/19940> holds various files of this Leiden University dissertation.

**Author:** Urban, Matthias

**Title:** Analyzability and semantic associations in referring expressions : a study in comparative lexicology

**Date:** 2012-10-10

# Appendix E

## Lexico-Semantic Associations

### INTRODUCTION AND NOTES ON DATA PRESENTATION

If motivated terms are coined or semantic extensions are institutionalized for whatever reasons, they necessarily bear some lexico-semantic associations by the very definition of the concept of lexical motivation, and there is little reason to believe that there is a correlation between the number of terms and their semantic structure.

Thus, however many motivated terms are found in an individual language, speakers in each case, for each meaning to be expressed, have to make a selection out of possible semantic associations. This section surveys the linguistic treatment of each meaning on the meaning list of the present study, and provides short accounts of the cross-linguistic variation found. In this sense, the model of this section is the still unrivalled work of Buck (1949) on Indo-European, and reference will be made throughout to it to compare the cross-linguistic findings with those found in this particular language family. However, there are also important differences from Buck's pioneering work: first, of course, the evidence presented here is based on a world-wide typological sample from languages of diverse genealogical affiliation, and is not restricted to one language family. Second, in contrast to Buck, the present account is synchronically oriented, as opposed to Buck's interweaving synchronic observations on lexico-semantic associations with diachronic developments. Third, the style of presentation is different, and the data provided here are more detailed than the discussion of the meanings in Buck: there will be ample data from the sampled languages to illustrate the observed patterns, and precise numbers as to their strength will be provided. In this sense, the chapter is a first step to answer the voices calling for a systematic cross-linguistic investigation of lexico-semantic associations discussed in chapter 2.

However, apart from a mere account of the data, there is more merit to such a discussion: Blank (2003) argues that Tagliavini's (1949) pioneering study of associations in terms for the 'pupil of the eye' allows to predict that it is highly likely that, if a complex term for this meaning should be coined anywhere, the associations will be drawn from the list compiled by Tagliavini. This, while at the same time extending such comparative overviews beyond a single concept, is another value of these data.

In order to make the presentation of the data maximally useful for further research and to make it readable independently from the framework developed in the present study, terminological peculiarities are kept to a minimum. However, sometimes it is useful to be able to have recourse to a number of terminological conventions to talk succinctly about the phenomena encountered. These require little to no additional theoretical assumptions, and thus do not distort the data in terms of a particular theoretical framework. SIMILARITY and CONTIGUITY are, purely descriptively, taken to be the fundamen-

tal semantic relations underlying METAPHOR and METONYMY, respectively. For instance, Cahuilla uses the same term for ‘finger’ and ‘hand’ (in more technical parlance that will be employed in the following, Cahuilla COLEXIFIES these meanings), while Aguaruna uses the same term for (colexifies) ‘finger’ and ‘branch.’ Now, the meanings ‘finger’ and ‘hand’ stand in a relationship of contiguity to each other (they are spatially adjacent to each other, or, since indeed the fingers are part of the hand, ‘finger’ is a meronym of ‘hand’), while this cannot be said of ‘finger’ and ‘branch’: rather, they are similar to each other in their longish shape and the fact that they protrude from a larger entity, the (rest of the) hand and the trunk of a tree respectively. In chapter 3, the semantic relationships of contiguity and similarity are defined in terms of test frames, and the categorization in the following relies on these as the criterion as to which fundamental semantic relation should be posited; however, since these test frames are explicitly designed to capture intuitions as to which relation is present it is not strictly speaking necessary for the reader to bear in mind in detail how these tests work in order to get an overview of the linguistic treatment of the concepts of interest to him/her. In analyzable terms which contain more than one lexical element, CONTIGUITY ANCHORING describes the semantics of one of the two constituents which ties back a metaphorical transfer accomplished by the other element to the semantic domain of the target concept. For instance, in Katcha *æ mǝ mbǝrǝ* ‘nostril,’ literally ‘eye of nose,’ there is a metaphorical transfer from *æ* ‘eye,’ the head of the complex term, to ‘nostril,’ while *mbǝrǝ* ‘nose’ is in a relation of contiguity to the target concept ‘nostril’ and anchors the metaphorical term in the domain of the target concept.

Some notes on presentation style are in order: each section begins with an overview of the most common lexico-semantic associations related to the meaning under discussion, in descending order of their occurrence in the languages of the sample. Further, there are percent values giving an idea of whether for this meaning, motivated terms are dominantly contiguity- or similarity-based (here, reported values often do not add up to the total number of motivated terms, which is either due to several possible analyses of the semantic association(s) or it being unclear), and whether they are typically by morphological analyzability or more often by colexification. Also, for each of the meanings, the number of languages for which an equivalent could be retrieved from the consulted sources is stated. In the body of the text, languages betraying a given recurrent lexico-semantic association are named first according to the macro area from Dryer (2005) they belong to, and are, within areas, listed alphabetically. Associations only found in one particular language of the sample are typically listed in a separate section that follows the one discussing recurring associations, except if they relate in some way to one of the associations discussed earlier in the text. If additional meanings mentioned in the consulted sources (other than ones that are so close to the item on the meaning list or one of the recurrent associations which is already discussed that mentioning them specifically seems redundant) are omitted from the discussion, this is indicated by phrases such as “*inter alia*.” This is mostly done for languages in which the relevant terms have a wide range of meanings none of which seem to be standing in any obvious semantic relation to one another. For this section in particular, it must be emphasized that formulations such as “the term *x* in language *y* also means *z*” or “colexifies *z*” does not necessarily entail the claim

that the meanings are related semantically. This is true generally, but should be borne in mind particularly when it comes to statements pertaining to Mandarin Chinese, where phonological changes discussed in § 5.4.2.3.2. lead to the collapse of a large number of erstwhile distinct lexical items.

When longer passages or longer glosses anywhere in the following are quoted directly from a consulted source, they are given in double quotation marks (this does not exclude the possibility that shorter glosses given in single quotes coincide with the gloss in the source). In the case that phonological processes alter the shape of the constituent morphemes of an analyzable term on the surface, the (inferred) morphological analysis is given in square brackets after the relevant object language term, as in Ket *destul* /dēs-d-ūl/, literally ‘eye-POSS-water’ and meaning ‘tear.’ Otherwise, the surface form in italics is segmented directly in order to save space. Grammatical material in glosses is, as in this example, printed in small caps and usually abbreviated; a list of these abbreviations can be found in the front matter. Possibly existing analyzable terms of the redundant type (as when the simplex *ange* in Kosarek Yale denotes both ‘umbilical cord’ and ‘navel,’ but there also is the complex term *ange lom*, with *lom* meaning ‘hole, valley,’ to single out the meaning ‘navel’) are usually not mentioned.

Discussions of individual meanings are referred to as “sections” along with their respective numbers in cross-references within this appendix, while cross-references to chapters in the main text are indicated by a paragraph sign (§).

The proxies, if any, that were accepted for a given meaning on the list are identical generally to those in Buck (1949), unless otherwise noted. For instance, as in Buck (1949), no attempt to distinguish between (bigger) ‘lake’ and (smaller) ‘pond’ is made. As noted in chapter three, a given term may exhibit more than one lexico-semantic association, either by virtue of being of the lexical type, or by being analyzable, but at the same time also colexifying more than one meaning. The latter fact in particular is a problem for a smooth and readable discussion of the associations found in terms for a given meaning, since it would require discussion constantly to jump back and forth between different associations. For instance, the Fijian word for ‘mirror,’ *i iloilo*, contains the instrument nominalizer *i* and a reduplication of the verb *ilo* ‘to look at.’ However, at the same time, it also colexifies ‘glass’ in general, a pattern that other languages exhibit as well, and thus Fijian has to be mentioned twice in the discussion, first in discussing terms derived from verbs meaning ‘to look,’ and second in terms colexifying ‘mirror’ with ‘glass.’ Since it is impossible to always do so without proliferating cross-references within each individual section ad absurdum, it is not always indicated which term bears which multiple semantic associations. This can be inferred simply from languages being mentioned twice, except if a language has several terms for the meaning and they exhibit different lexico-semantic associations. Conversely, however, it may be that languages have synonyms for the meanings in question and different ones bear different associations. That is, if a language is mentioned several times in each section, this does not necessarily always entail that it is always the same term bearing all associations.

Moreover, not all patterns of colexification are mentioned for constituents in complex terms. For instance, terms for ‘tear’ often consist of elements meaning ‘eye’ and

‘water,’ with ‘eye’ having secondary readings (see section 140 and § 6.2.3.1.) which are not mentioned since clearly it is the core meaning ‘eye’ that is relevant in the conceptualization of ‘tear.’ Any other way of proceeding would yield a resulting discussion that is highly cumbersome to read. This policy is departed from in cases where it seems unclear which sense of the constituent is relevant for the conceptualization of the complex term, in which case different meanings of constituents are separated by slashes in the gloss.

A further issue is ambiguities arising due to strictly speaking conflicting information in two-way dictionaries. For instance, Nez Perce *hími* is stated in the Nez Perce-English section of the consulted source to mean ‘mouth; mouth of river, cave,’ but when one seeks to identify the Nez Perce term for ‘lip’ from the English-Nez Perce section, one also encounters *hími* as one of the equivalents. In this and other such cases, all information was taken into account, that is, Nez Perce is mentioned as a language colexifying ‘lip’ and ‘mouth’ and is mentioned as such in both relevant sections, in spite of being not explicitly glossed as ‘lip’ in the dictionary.

### 1. *The Animal*

Representation: 83%

Motivated: 43.1%

Thereof Analyzable: 26.8%

Thereof Colexifying: 18.3%

Thereof by Contiguity: 32.8%

Thereof by Similarity: 1.6%

Recurrent associated meanings: thing, meat, life/be alive, bird, livestock, cattle, move, pig, dog, insect, forest/wood, kill, move on ground, brute, person, land

A common association, by contiguity, is that with ‘meat’ or ‘flesh,’ since obviously animals are the source of meat to be consumed by humans. The association is mostly realized by colexification and occurs in Efik, Hausa, Ngambay, Noni, Yoruba, Gurindji, Yir Yoront, Abzakh Adyghe, Sora, and Yanomámi, and by noun class alternation in Swahili.

Another common pattern is to have terms for ‘animal’ which also at the same time mean ‘thing’ very generally. This is found by colexification in Buli, Kwoma, Arabela, Chayahuita, Lengua, Wayampi (where the relevant term can also mean ‘kind of’ and ‘luggage’), and Samoan. In Lengua, as well as in Khalkha, the term may also refer to a ‘person’ and in Buli also to a ‘figure, unrecognizable person;’ compare also Copainalá Zoque *copan* ‘animal’ and *pan* ‘person.’ Furthermore, in Katcha, there are the complex terms *nimo mo tile* ‘thing of forest’ and *nimo mo di* ‘thing of house’ to refer to wild as opposed to domestic animals respectively (for the former term, compare also Bora *bájú-e-jpi* ‘mountain/forest-belong.to-CL.M.S.’ and Gravelle’s 2004: 375 statement that a similar term is found in Meyah; the distinction between domestic and wild animals is also made in Rendille, Oneida, and Aymara; for Cushitic specifically see also discussion in Sasse 2002). Similarly, Itzaj has *b’a’al-che* ‘thing-wood/tree’ for ‘animal, fauna.’ Still further, Bwe Karen has the somewhat unclear term *de-pho=de-wə* ‘thing-child/little.one=thing-bug’ glossed as ‘dumb creatures, animals’ (compare colexification of ‘animal’ and ‘bug’ in Highland Chontal and of ‘animal’ and ‘insect’ in Nivkh, Kildin Saami, and Hawaiian). There is a semianalyzable term with the identifiable constituent meaning ‘thing’ in Sko.

Mandarin has *dong4-wu4* ‘move-object/being,’ and Japanese and Vietnamese, due to Chinese influence, *dō-butsu* ‘move-thing’ and *động vật* ‘move object’ respectively. Moreover, Hani, a language also spoken in Southeast Asia, has *nivzeig* ~ *nilzeig*, maybe related to *niv* ‘moving; action’ (though note also *nil*, meaning ‘the outside, wild’ inter alia), Nuuchahnulth has *saštuup* ~ *saxtuup* /sa-tuʔp/ ‘crawl.on.all.fours-creature,’ (with *tuʔp* also glossed as ‘thing’ generally as well as ‘kind, sort,’ compare Chukchi *yennik* which is perhaps related to -*ɲərtə*- ‘four’). Similarly, Blackfoot has *iksowáʔpomaahkaa* /iksow-áʔp-omááhkaa/ ‘at.ground.level-about/around-move.along.on.foot,’ and Lakhota *wamákhaská* is literally translatable as “those moving about on earth.”

Returning to associations with ‘thing,’ Khoekhoe has *ûitsama tsaba xūn*, with *ûitsama* meaning ‘living’ (derived from *ûi* ‘to escape, to escape death, be alive’), and *xūn* referring to ‘possessions’ and ‘livestock.’ Both associations recur: ‘livestock’ is also colexified in Rendille, Basque, and Yay, and relevant terms in Gurindji, Bezhta, and Samoan also denote ‘cattle’ specifically. As for the association with ‘life’ or ‘be alive,’ there are derived terms in Muna and perhaps in Greek, Kiliwa has *t-kw+ipaa-y* ‘OBJ-WH+be.alive-ATT,’ Nez Perce *waqíswitin* /waqíswit-i-n/ ‘life-with,’ and Carrier *rhenna* ‘large animal’ is a verbal noun derived from *rhesna* ‘to be alive.’ Lenakel combines the association with the already familiar pattern relating ‘animal’ and ‘thing:’ *nar amíuh* is analyzable as ‘thing alive,’ and Piro *giwekachri yotaljetachri* relates again to the association with ‘movement’: it contains elements meaning ‘to be alive’ and ‘to move.’ Yoruba *ẹlẹmì* is analyzable as /ẹlẹ-ẹmì/ ‘owner-breath/life’ (there is also the term *ẹ-dá* ‘NMLZ-to.create’ for ‘creature, creation’). Further, Central Yup’ik *ungungssiq* contains the base *unguva* ‘life,’ in Samoan, there is the redundant analyzable term *meaola*, with *mea* meaning ‘thing, animal’ (and also ‘genitals’ and ‘place’ generally inter alia), and *ola* ‘living,’ and there are (probably) further terms that are diachronically related to verbs meaning ‘to live’ in Khalkha (here, the relevant term seems to have further connections to terms related to ‘breathing’), and Kolya Yukaghir.

Upper Chehalis *xəs=áy=tmʃ* is (semi)analyzable as ‘bad=??=land,’ and Central Yup’ik has a term for land animals, *nuna-miutaq*, which is analyzable as ‘land-one.whose.proper.place.is.’ A pattern peculiar to New Guinea is to have dvandva compounds expressing the concept ‘animal,’ with the constituents denoting particular animal species: thus, Kyaka has *suwua-pe saa-pe mena-pe pyasingi* ‘dog-ASSOC furred.animals-ASSOC pig-ASSOC mixed.assorted.group,’ Sentani *obo-joku* ‘pig-dog,’ and Takia *bor-goun* likewise ‘pig-dog,’ but here the compound has reference to ‘domestic animals’ specifically. Hawaiian has *holoholo-na* ‘walk-NMLZ,’ and there is a semianalyzable term featuring a constituent meaning ‘walk’ in Great Andamanese. Haida has *gina tiʔaraa* ‘creature/thing be.killable’ for ‘land mammals’ specifically, and the Oneida term for ‘wild animals,’ *kutilyoʔshúha*, is likewise derived from a verb meaning ‘to beat, kill.’ Lesser Antillean Creole French and Basque colexify ‘animal’ with ‘brute.’ Kwoma *boboy* also means ‘plant,’ and the association might also be present in Abipón, but this is considered unsure in the consulted source.

Berik and Biloxi make a distinction between ‘female animal’ and ‘male animal’ (the Berik term for ‘wild animal,’ *giri*, also means ‘deep river’). Upper Chehalis, Highland Chontal, Huambisa, Jarawara, Fijian, Kapingamarangi, Rotuman, and Samoan colexify

‘animal’ with ‘bird,’ that is, a lower rank in ethnobiological taxonomies, and ‘animal’ is colexified with specific animals on the generic level in Dadibi (‘cuscus’), Sora (‘pig’), Koly-ma Yukaghir (‘elk’), Pawnee (‘deer’), and Central Yup’ik (‘bear’).

Other associations include: Buli *dung* also means ‘to press down, press out’ inter alia, and Efik *u’nam* might be related to *nam* with the basic meaning ‘to do, make’ (considered unsure in the consulted source). Ngambay *da* also means “to assemble something” and also denotes a “kind of tree, bark strings to attach the roof of a house with.” Noni *nyam* might consist of the verb *yam* ‘to suck’ and a prefixal noun class marker. The Anggor term *nine-hondi* is analyzable as ‘fur/feather-mother.’ Khalkha *aduyusu(n)* is derived from *aduyu(n)* ‘herd of horses, horse’ by means of the suffix *-sun* the function of which is, according to Poppe (1954: 44), “to form nouns of which the meaning is usually the same as that of the primary word.” Ket has *assel* /ēs-sèl/ ‘wild reindeer,’ which is likely an instance of a so-called markedness reversal (Berlin 1972, Witkowski and Brown 1983). The Cahuilla term *ʔiʔihɪɪvəʃ* is likely to contain *-ʔi* ‘leg, foot, footstep, track.’ Kashaya *ʃiʔbaʃi* contains *ʃiʔba* ‘body,’ and Kiowa has a term for domestic animal, *ɣɪtɔɪ-dou*, which is analyzable as ‘go.live.with-have.’ Quileute *ʔixʷáʔtso* appears to be related lexically to *ʔixʷáʔtsil* ‘to hunt,’ and Tuscarora *yuʔtikɛhrat* contains the roots *-(ɛ)ʔtikɛh(r)-* ‘mind’ and *-aT-* ‘stand.’ The Yana term *mooyau(na)* consists of *ma-* ‘to eat’ and the nominalizer *-yau(na)*, and Yuki *heʔlikʔke* appears to contain *lik* ‘to swallow.’ Embera *ārimārā* also means ‘inhuman’ and ‘cruel,’ Sáliba *omaĩdi* also ‘heart,’ and Yanomámi *yaro* can in some contexts also refer to an ‘enemy warrior.’ Finally, Kapingamarangi *manu* is also the name of a constellation involving the star Sirius.

## 2. The Ashes

Representation: 96%

Motivated: 38.5%

Thereof Analyzable: 17.4%                      Thereof Colexifying: 21.0%

Thereof by Contiguity: 5.3%                      Thereof by Similarity: 18.4%

Recurrent associated meanings: dust, fire, powder, embers, coal, faeces, dirt, gray,  
kitchen/fireplace, wood, lime, sand, black, soot, blue, soil, flour, burn, feather

Intra-domain associations with other fire-related terms, namely ‘coal’ and ‘embers,’ are frequent for ‘ashes’ (or ‘cinders’), see also sections 13 and 19. Five sampled languages, Burarra, Kwoma, Yir Yoront, Khalkha, and Sedang colexify ‘ashes’ and ‘coal,’ and there are derived terms in Wintu (colexifying ‘soot, soot carried by the wind,’ for the association with ‘soot,’ compare also Hawaiian *paʻu ahi* ‘soot fire,’ denoting ‘soot’ next to “black cinder sand or ash”) and Great Andamanese. In a parallel fashion, Kosarek Yale, Khalkha, Carrier, Kiliwa (by the complex term *kw+pal* ‘PERF+hot’), Cubeo, and Ancash Quechua colexify ‘embers,’ while Aymara has *nina sank’a* and *nina japu* ‘fire/embers embers.’ More generally, mirroring an association also diachronically attested in Indo-European (Buck 1949: 73-74), many sampled languages employ the same term for ‘ashes’ and ‘dust,’ namely Efik (where the relevant terms also can refer to a “dimness of vision as if a mist were before the eyes” and a kind of spearmint), Berik, Burarra (where the relevant term more



precisely means “little particles, as ashes and sand mixed where a fire has been burning” as well as ‘dust’), Lavukaleve, Mali, Nunggubuyu, Rotokas (colexifying also ‘wood chips’), Basque, Upper Chehalis (‘cold ash’ specifically), Chukchi, Biloxi, Carrier, Highland Chontal, Lesser Antillean Creole French, Lake Miwok, Oneida, Tuscarora, Wintu (where *bukul* contains *buk* ‘dark’), Yuki (where *poʔoʔel* contains *poʔ-* ‘burn’ and *poʔ-* ‘gray’), Cavineña (where the relevant term means ‘dust cloud’ more specifically), Hani, and Bwe Karen. Alternatively, some languages have complex terms for ‘ashes’ based on ‘dust,’ mostly with ‘fire’ acting as a contiguity anchor, as in Northern Yana *tabʔlaawi-ʔau(na)* ‘dust-fire.’ Such terms are also found in Carrier, Nez Perce, and Wichí; alternatively, in Pawnee, ‘ashes’ is *karaak-tuuhcu*?, analyzable as /itkaar-haak-tuuhc-uʔ/ ‘dust-wood-??-NOM’ and in Hawaiian *lepo uli*, analyzable as ‘dust/dirt/soil/rubbish/excrement dark.color’ (a semianalyzable term with the constituent ‘dust’ is found in Kashaya and San Lucas Quiaviní Zapotec).

For the Yuki association with ‘gray,’ note also that Muna *ghabu* may also refer to the “grey colour of certain fruits (indicating old age)” (cf. also *ghabu* ~ *ghadu* ‘enormous, huge’), as well as that Tuscarora *uʔkəhreh* colexifies ‘ashes,’ ‘dust,’ and ‘gray,’ Huambisa ‘ashes,’ ‘gray,’ and ‘lead,’ and Hawaiian *lehu* means ‘gray,’ but is used more specifically to describe the color of chickens inter alia. On a related note, Yanomámi *ushi pë* appears to consist of *ushi* ‘dark blue, dark violet, bruise, ripe’ and the quantal classifier *pë* (see § 4.4.1.1.), while Yay *taw*<sup>5</sup> colexifies ‘ashes’ with “blue, ash-color.” Colexification with ‘dirt’ is attested for Oneida and Wintu, here by the same semianalyzable term mentioned above (in both languages, ‘soil’ is also colexified, and Yanomámi colexifies a specific type of dark soil with ‘ashes’), and complex terms betraying this semantic association are found in Yir Yoront (*thum-nhur* ‘fire-dirt’), Wappo (*hél-pi-pol* ‘fire-from-dirt’) and in Hawaiian, as seen above. Still more generally, colexification with ‘powder’ occurs in Ngambay, Kyaka (‘white powder’ specifically), Basque, Chukchi, Cheyenne, Highland Chontal, and Bora, and a realization of the association by morphological complexity is found in Piro (*tšitšj-pahi* ‘hearth-fire/firewood-powder’) and Manange (*4mje-pʰra* ‘fire-powder’), while there is a semianalyzable term in San Lucas Quiaviní Zapotec. More specifically, Hausa colexifies ‘ashes’ with ‘pounded flour’ in particular (compare Bororo *joru-gudu* ~ *ru-gudu* ‘fire-flour’). These semantic associations mirror closely those reported for ‘dust’ in section 16 (see also Buck 1949: 74 for Indo-European evidence).

Metaphor-driven associations linked with ‘ashes’ specifically, but usually not with ‘dirt’ (with the exception of the evidence from Hawaiian already discussed above) are that with ‘faeces’ or ‘excrement,’ occurring by colexification in Kwoma (the term also colexifies ‘flower’) and by analyzable terms of the structure ‘fire-faeces’ in Cayapa, Bislama, and Takia, while Itzaj has *taʼan* /taʼ-Vn/ ‘excrement-DERIV’ (note also the somewhat unclear case of Kiowa *sāʼ-pʼh̄n*, perhaps ‘excrement-cloud/sky’). The Itzaj term betraying this connection in addition colexifies also ‘lime,’ while in Copainalá Zoque *cuy-jam* is analyzable as ‘tree/wood-lime.’ Outside Mesoamerica, Tetun has *ahu-metan* ‘lime-black’ (for which in turn compare Santiago Mexquititlan Otomí *ʼbo-spi* ‘black-fire’ and Kyaka *keyihapa now*, where *keyihapa* is ‘black’ and *now* is a term for color and coloring earth pigments), and Fijian colexifies ‘ashes’ with ‘slacked lime’ specifically (the relevant term *dravu* means ‘to rub the head with ashes’ as a verb). There is an association with ‘feather’ found in Tsafiki

(*nin fu* ‘fire feather/body.hair’) and Great Andamanese (*châpa-l’ig-pîd* ‘firewood-??-feather’).

Further complex terms in which ‘fire’ acts as a contiguity anchor include Lesser Antillean Creole French *sann dife* ‘sand fire’ (compare colexification of ‘ashes’ and ‘sand’ in Cayapa and Hawaiian), and Sedang *pló on* ‘fermenting.agent fire.’ Semianalyzable terms with ‘fire’ are found in addition in Mbum, Toaripi, Kosarek Yale, Nez Perce, and Pipil. Sedang *pló on* ‘fermenting.agent fire’ also colexifies ‘hearth’ or ‘fireplace,’ and the association with ‘fireplace’ or ‘kitchen’ is mirrored by unanalyzable or semianalyzable terms in Muna, Sahu, and Toaripi (note also the apparent relationship Guaraní *tanimbu* ‘ashes, remains of things that were, debris’ and *tanimbupa* ‘kitchen, hearth’). There are also some terms which highlight that the ashes are the remnants of a fire. These include Abzakh Adyghe *safe* /sə-efe ~ sə-afe/ ‘burn-REST’ (compare the association between ‘ashes’ and verbs meaning ‘to burn’ in Indo-European languages, Buck 1949: 73), Central Yup’ik *qamlleq* /qame-lleq/ ‘die.down-one.that.has,’ and Samoan *lefulefu*, where the reduplication base *lefu* means ‘for a fire to go out’ (compare again Kiliwa *kw+pal* ‘PERF-hot’ as well as Welsh *gweddillion* ‘ashes,’ which is related to *gweddill* ‘remnant’).

Further associations include: Hausa *ru’bushi* ‘hot ashes’ is also the name of a kind of sweet pastry inter alia, and *raushi* ‘hot fine ash’ also means, inter alia, ‘softness, tenderness.’ Kyaka *pee* also is a general term for a ‘receptable’ and specific receptables inter alia, and Chukchi *piŋ(pij)* also means ‘falling snow.’ Japanese *hai* also means ‘yes’ (with the term in the meaning ‘ash’ perhaps being borrowed from Chinese). Kwoma has *hikishebo*, colexifying ‘ash’ with “black paint, black pigment used to make paint” and “object burnt by fire” as well as “earth blackened by fire.” Ngaanyatjarra *tjurnpa* also means ‘husks,’ and Rotokas *gavuta* is also glossed as ‘bed of fire.’ Badaga *budi* also means ‘ashmound’ as well as ‘contrary,’ Khalkha *coy* also ‘sparks’ and, figuratively, ‘glory’ and ‘energy’ inter alia, and Welsh *ulw* also ‘utterly.’ Upper Chehalis *sqʷəlŋš* is derived from *qʷəl* “roast, cook, ripe, ripen” and contains the lexical affix =*nš* ‘basket’ (or ‘ball’). Tuscarora *uhséhareh* means also ‘lye,’ and the first variant of Abipón *-aci* is identical formally with terms for ‘tongue’ and ‘tear.’ Cashinahua *mapu* colexifies ‘ashes’ with ‘brain’ (the tertium comparationis being the gray color of both?), ‘soap,’ and ‘handle of a tool’ inter alia. Jarawara *hasawiri/hasawiri* also means ‘smoke,’ and Kapingamarangi *lehu* also ‘smegma.’<sup>1</sup> Sedang *trôi* is also the name of a ‘leaf-eating ant,’ and Bislama *asis faea* can also refer to ‘relic, remnant’ in non-standard usage.

### 3. The Bark

Representation: 93%

Motivated: 68%

Thereof Analyzable: 23.7%

Thereof Colexifying: 44.6%

Thereof by Contiguity: 0.6%

Thereof by Similarity: 67%

Recurrent associated meanings: skin, peel/rind/shell, tree/wood, hide/leather, cover, scale, husk/chaff, crust, clothing, pod, scab of wound, hard, tire, fingernail

<sup>1</sup> “smegna” in the source.

Clearly, the most frequent association is that with 'skin.' It occurs either by colexification or analyzable terms (with 'tree' or, less frequently, 'wood' as contiguity anchor, as in Yei *pər par* 'tree skin' or Abzakh Adyghe *pxa-sʰe* 'wood-skin,' and sometimes also, due to the colexification of 'tree' and 'wood' in many languages as reported in section 65, with both, as in Berik *ti tifi* 'tree/wood skin'). Such terms are attested in as many as 82 sampled languages. 55 languages are of the colexifying type: Bakueri, Efik, Yoruba, Anggor, Buin, Burarra (where the term also colexifies 'wrapper, outer case' and is perhaps related to a word for 'clan'), Gurindji, Kwoma, Kyaka, Lavukaleve (colexifying 'bark of coconut tree' specifically with 'skin'), Mali (colexifying also 'body'), Muna, Ngaanyatjarra, Nunggubuyu, Sahu, Sentani (where the relevant term is also the name of a palm species and a discourse particle), Sko, Southeastern and Western Tasmanian, Toaripi, Kosarek Yale, Basque, Bezhta, Biloxi, Chickasaw, Ineseño Chumash (where the relevant term means 'smooth bark' specifically), Comanche, Haida, Pipil, Xicotepec de Juárez Totonac ('skin of leg of a person' specifically), Yaqui, Copainalá Zoque, Abipón, Bororo, Carib, Cashinahua, Cayapa, Guaraní, Hupda (colexifying also 'dish, plate, food'), Jarawara, Kaingang, Maxakalí, Piro, Ancash Quechua, Tsafiki, Wayampi, Yanomámi, Bislama, Fijian, Hawaiian, Bwe Karen, Lenakel, Malagasy, White Hmong, Rotuman, and Samoan. There is a derived term in Great Andamanese, and twenty-four languages have complex terms with 'tree' and/or 'wood' acting as contiguity anchor: Mbum, Ngambay (*ngóy gírì kake* 'skin/peanut.shell/peeling behind tree'), Baruya, Berik, Kaluli, Meyah, One, Waris (where 'skin' is colexified with 'blood'), Yei, Yir Yoront, Abzakh Adyghe, Japanese, Kolyma Yukaghir, Wappo, Yuki, Cavineña, Miskito, Sáliba, Rama, Hani, Mandarin, Takia, Tetun (where 'tree' and 'plant' are colexified), and Yay; further, Noni has a term with this structure not for 'bark' generally, but for 'skinned bark' specifically. The association is diachronically detectable also in Chukchi. Given the fact that there are 138 languages in the sample for which data are available, this means that the association occurs in almost 60 per cent of the sampled languages, and that in all regions of the world. Kiowa is a little different in that here *ʔæ'kʷə* seems to be a diminutive of *k'æ* 'skin' indicated by vowel nasalization combined with an unknown second element, and Ket *iŋ-ol-t* is analyzable as 'skin.PL-covering-NMLZ.' Furthermore, Cubeo *jocw-curi* is analyzable as 'wood-CLASS.TEXTURED.COVER,' and in Yanomámi *hisi*, *hi* is a term for 'wood' which also acts as a classifier for plants with wooden trunk and *si* means 'cover' generally; there also is an alternative reduplicated term *sisi* for 'bark.' Note also that there are many languages with redundantly complex terms, such as Dongolese Nubian *ǰōwwi-ŋ-gábad* 'tree-GEN-bark.' Buli also colexifies 'bark' and "hard skin" (e.g. of a crocodile)" specifically, and there are semianalyzable terms where the identifiable constituent means 'skin' in Copainalá Zoque (where the unknown element recurs in many terms for specific trees) and Great Andamanese.

Due to colexification of 'skin' with 'leather' or 'hide' (compare section 135) a subset of these languages also betrays this association by colexification. This is the case in Kyaka, Muna, Sentani, Ket (by the analyzable term *iŋ-ol-t* 'skin.PL-covering-NMLZ'), Cahuilla, Pipil, Yaqui, Copainalá Zoque, Abipón, Bororo, Carib, Cashinahua, Guaraní, Jarawara, Piro, Ancash Quechua, Tsafiki, Yanomámi, Hawaiian, Bwe Karen, White Hmong, Rotuman, and Samoan. Complex terms with the 'tree/wood' contiguity anchor are found in Highland

Chontal, Pawnee (where the constituent denoting ‘hide’ also means ‘peeling’ generally inter alia), Wappo, Cavineña, Miskito, Sáliba, Mandarin, and Tetun. In addition, in Sedang, *kotôu* colexifies “bark; peeling, rind, husk; egg shell; outer edge of ...,” that is, a different term for ‘skin’ is used, and the same is true of Chayahuita.

Another quite frequent association also exhibited by the Sedang term is that with the meaning ‘shell’ in general and/or ‘rind, peel,’ as for instance of a fruit, an egg, or a mussel more specifically, occurring in fifty-nine sampled languages, by colexification in Bakueri, Buli, Hausa, Yoruba, Kyaka, Ngaanyatjarra, Mali, Muna, Abzakh Adyghe, Badaga, Greek, Khalkha (also colexifying ‘bark of birch tree’ specifically), Sora, Kildin Saami (here the relevant term *kērr* also means ‘hard, rough,’ as is the case in White Hmong; similarly, Hawaiian has a derived term, which is, however, rare), Cahuilla, Chickasaw, Ineseño Chumash, Haida, Kiliwa, Lakota, Lesser Antillean Creole French, San Mateo del Mar Huave, Nez Perce, Pipil, Central Yup’ik, Copainalá Zoque, Aguaruna, Bororo, Cashinahua, Cayapa, Guaraní, Huambisa, Hupda, Jarawara, Kaingang, Macaguán, Maxakalí, Piro, Ancash Quechua, Tsafiki, Yanomámi, Bislama, Fijian, Hawaiian (where the term has a very broad range of meanings, including also ‘scalp’ and ‘surface, area’ inter alia), Rotuman, Sedang, Tetun, and Vietnamese, and by complex terms in Efik (*ikpök ètü* ‘skin/bark/scale/shell/husk/pod tree’), Ngambay (*ngóy giri kake* ‘skin/peanut.shell/peeling behind tree’), San Mateo del Mar Huave (*mipang xiül* ‘shell tree,’ this is a different term from the colexifying one mentioned above), Itzaj (*pach che* ‘shell tree’), Yuki (*ʔol šil* ‘tree skin/shell’), Copainalá Zoque (*cu’y-unaca* ‘tree-skin/shell’), Arabela (*naana co* ‘tree shell’), Bora (*kó-mi:ʔo* ‘wood-SCM.hard.shell’ and *úme-he-mi:ʔo* ‘tree-CL.tree-SCM.hard.shell’), Rama (*katuuk /kát-úk/* ‘tree-skin/coat/rind’), Sáliba, where *ixexe* colexifies ‘skin,’ ‘hide,’ and ‘shell,’ while ‘bark’ is *nugu ixexe* (*nugu*, ‘trees’), and Hani (*albol alhov* ‘tree-skin/peel/outer.layer.of.grain’); moreover, Bezhta *yicalo* contains an element meaning ‘peel.’ In most of these languages, ‘skin’ is also colexified; exceptions are Kiliwa, where *yal* is only used for ‘bark’ and ‘shell,’ but not for ‘skin,’ Tetun, where *kakun* colexifies ‘peel,’ ‘shell’ and ‘bark,’ but not ‘skin,’ Pawnee, which has *raakickuusu*?, analyzable as /raak-(i)ckuus-u’/ ‘tree/wood-peeling-NOM’ (with *(i)ckuus*-also being capable of referring to ‘hide, shell’ and other similar meanings, but apparently not to ‘human skin’), as well as Bezhta, where the term already mentioned is derived from an element meaning ‘peel.’ As for the association with ‘tree, wood,’ there is one language, Abipón, which colexifies ‘bark’ and ‘tree’ directly.

Furthermore, nine sampled languages colexify ‘bark’ with ‘scale,’ namely Hausa, Lavukaleve (where the relevant term also denotes the “place coconut stalk hangs from”), Bezhta, Biloxi (where the relevant term also colexifies ‘horns,’ ‘hoofs’ and ‘nails on fingers and toes,’ which latter association is also present in an analyzable term in Rama), Nez Perce, Central Yup’ik, Guaraní, Jarawara, and Bwe Karen, and Efik has the complex term already mentioned. In most of the above languages, the relevant terms have a rather broad semantics which, from an extensional point of view, colexify at least one other, but frequently many more, meanings figuring in this report, so that their semantics, in these cases, may also be described intensionally as ‘outer surface’ of anything. In fact, in Sko, Basque, Chickasaw, Comanche, Haida, Lakota, Guaraní, Huambisa, Maxakalí, Piro, Yanomámi, Lenakel, and Bislama, the relevant terms are also explicitly glossed as meaning

‘cover, covering’ (in Haida also as ‘surface of waves’ and in Lakhotá also as ‘envelopes’ and ‘wrappings’ specifically). Instructive in this context is the example given in the consulted source for Kaingang *fár*: *kar vỹ, fár nĩ: ka kar, nén tánh kar, êg mêng kar, êg ke gé* “everything has *fár*: trees, plants, animals, we do, too.”<sup>2</sup> However, Bezhta *qal* exclusively denotes ‘scale’ and ‘bark,’ while for ‘skin’ and other meanings alternative terms are available.

Very broad semantic range of the relevant terms is also the case for the terms in which the following associations figure, unless otherwise noted as for Bezhta above. In seven sampled languages, Yoruba, Bakueri, Anggor, Muna, Sahu, Bislama, and Sedang, ‘bark’ and ‘husk’ or ‘chaff’ are colexified (in Anggor also ‘container’), while Efik has the complex terms already mentioned. Kyaka, Muna, Basque, Greek, Sora, Hawaiian, and Rotuman colexify ‘crust’ (Greek ‘crust of earth’ and Sora ‘incrustation’ specifically; the relevant Hawaiian term is that derived from ‘hard’ mentioned above and may also mean “cooked crisp, as pig” inter alia), and in four sampled languages, Anggor, Sko, San Lucas Quiaviní Zapotec, and Piro, ‘bark’ and ‘cloth,’ ‘clothes,’ or ‘clothing’ are colexified (in San Lucas Quiaviní Zapotec also ‘rag used as diapers’ is colexified, while the relevant Sko term also denotes a tree species). Abzakh Adyghe and Yoruba also colexify ‘pod’ (‘pod’ is colexified in the element meaning ‘skin’ and other things figuring in the complex Efik term mentioned above), while Buli, Hausa and Hawaiian in addition colexify ‘bark’ and ‘scab of wound’ (Hausa also “scurf of a scalp disease” and “bits of skin from desquamation”). Rotuman and Samoan also colexify ‘(rubber) tire’ (Samoan in addition “kin, kinsmen” as well as “belt (for a machine)” and ‘foreskin’).

Other infrequent lexico-semantic patterns include: Buli *tapagi* also colexifies “board, flat piece of wood,” while Khoekhoe *tammi* is derived from the verb *tam*- ‘debark.’ Rendille has unrelated terms for “soft inner tree bark” and “soft outer tree bark,” with the former also meaning “soft membranes” and “bags of waters in pregnancy” of animals, while in Nivkh, *oym* also means ‘root.’ Burarra *mun-ngarnama* is analyzable as ‘CLASS.DOMESTIC-inner.thigh,’ which illustrates the widespread mapping of body-part terms onto entities of the physical world in Australia, and in Kyaka *imbu* ‘hard shell, bark’ may also be used to refer to “anger, annoyance” as well as a ‘bowstring.’ Muna *kuli* also means to “have a simple meal without side dish.” Ngaanyatjarra *miri* may also refer to “people of a certain skin colour, race of people” as well as “get a shock, be shocked,” *pangki* (Northern dialect) experienced semantic extension to ‘orange’ by synecdoche and *piilyurru* also denotes the ‘carpet snake, woma python,’ “the shed skin of a snake or goanna,” and a “chrysalis case, shell enclosing a grub before it turns into a moth” (see Evans 1992 on such associations in languages of Australia). Sko *ró* also denotes a tree species, and Basque colexifies “exterior, outside” and ‘surface’ generally, as well as ‘cover,’ ‘spread,’ ‘case,’ ‘sheet,’ and ‘rascal.’ Chukchi colexifies ‘bark’ with ‘edible whale skin,’ and Khalkha *xoltusu(n) ~ xoltasu(n)* also denotes ‘monkshood’ (*aconitum variegatum*). Xicotepec de Juárez Totonac colexifies ‘copal.’ The Wintu term *laplah* contains *lap* ‘lying down,’ *čahay* also means ‘sheath’ generally and denotes the ‘inner bark of maple’ as well as “deerleaf, cowleaf, the fat sheath veiling of the stomach of an animal ...” specifically, and *qoʔq* also

<sup>2</sup> Original Portuguese translation from the source: “Tudo tem ‘pele’: árvores, plantas, animais, nós também.”

means 'to pull of chunks.' Abipón has *l-aoel-ag-Ra* 'POSS.INDEF/3SG-heart/guts-PRODUCT-ABSTR.' Guaraní *ape* colexifies 'back' and *pire* 'movie.' Maxakalí *xax* also means to "seek, hunt, long for." Piro *mta* may also refer to a 'surface' or 'mat,' the Rama term *sarpan* also to 'bast,' Yanomámi *hisi* also to a 'wasp,' and Wayampi *pilɛ* also to any sort of covering of artifacts. Yanomámi colexifies 'bark' with 'bud,' Bwe Karen with "to scratch, claw, maul," Lenakel with 'cover of book' (the term is said to be capable of referring to "any ... outside covering" in the source), and finally, Bislama *skin* is also used with the meanings "very strong, effective, very good, cool."

#### 4. The Bay

Representation: 32%

Motivated: 90.3%

Thereof Analyzable: 47.3%

Thereof Colexifying: 44.8%

Thereof by Contiguity: 69.2%

Thereof by Similarity: 13.9%

Recurrent associated meanings: corner, sea, harbor, lake, valley, crooked/crookedness, elbow, mouth, breast, inside/interior, bend of river

Five sampled languages -and this already is the most frequently occurring association-colexify 'bay' (or 'inlet,' 'gulf' which were accepted as proxies) with 'corner' more generally. These are Nunggubuyu, Meyah, Khalkha, Toba, and Hawaiian (where the relevant term may also refer to a 'cove,' 'indentation' and 'cell of beehive' inter alia). Tetun has the complex terms *tasi-lidun* 'sea-corner' and *tasi-sikun* 'sea-elbow/corner' (in addition, the language also has the terms *tasi-keta* 'sea-to.separate,' *tasi-nanál* 'sea-tongue' and *tasi-soman* 'sea-companion;' for the second term mentioned, compare colexification of 'elbow' and 'cubit' in Khalkha). For the association with 'corner,' compare that with 'curved shape' reported by Buck (1949: 39) for Indo-European and the somewhat similar analyzable terms *kjamtsə-kʰa* 'ocean-edge' in Manange, colexification of 'to bend' and 'crooked, winding' in Mandarin, and *helodranomàsina* /hèloka-ranomàsina/ 'crookedness-sea' in Malagasy. In four sampled languages, terms for 'bay' are lexically connected with the meaning 'lake,' by colexification in Nez Perce, Santiago Mexquititlan Otomí and Rotuman (in the latter language by the analyzable term *tən hăe* /tənu hăe/ 'water contain;' the term can also refer to a 'puddle'), while Berik has the complex term *fo boga* /fo bogana/ 'river/lake calm.' Again four sampled languages, Burarra, Lavukaleve, Lesser Antillean Creole French, and Lenakel colexify 'bay' with 'harbor' (compare again Buck 1949: 39 for evidence from Irish and Germanic); in Burarra the relevant term is semianalyzable and contains a verb meaning 'to be helpless.' Another term for 'bay' in which the semantics of the constituents suggest that 'harbor' is a secondary meaning is Miskito *slaup yukuwaik* 'boat shelter.' Lenakel *nouanhulin* is analyzable as /noua-nhul-n/ 'opening-mouth-3SG.POSS.' For this term, compare also Tehuelche *k'on k'ork'en* 'bay, shore of river' which may alongside *k'on ~ k'on* 'river' contain *k'onk'en ~ konk'en ~ konken* 'mouth' (or perhaps *k'o:r ~ k'or ~ k'o?r* 'pasture?'), as well as the common 'mouth'-metaphor in terms for estuary reported in section 20. Two sampled languages of Europe, Basque and Greek, colexify 'breast' and 'bosom' (Basque also 'stomach, guts' and Greek also "apopleptic fit"); this is a pattern peculiar to Europe (see Buck 1949:

38-39 for further Indo-European evidence and § 6.4.3.3. for some details on the history), and in three other sampled languages, Central Yup'ik, Hawaiian, and Takia, associations with 'valley' or 'depression' are found: Takia *ilo-* colexifies "inside, emotions, valley, bay," Hawaiian *Hono-* colexifies 'bay, gulch, valley' (but occurs only as part of place names), and in Central Yup'ik (Yukon dialect), *ilutak* contains the base *ilu-* 'interior' and colexifies 'valley, dip.' Somewhat similarly, *ilutuqaq* /*ilutu-tuqaq*/ 'be.deep-one.that.has.to.a.large.degree' colexifies 'bay' with 'depression' and 'deep place' (there are also the terms *kangiqaq* /*kangiq-qaq*/ 'headwaters-area.of.possessor' and *kangir-rluk* 'headwaters-one.that.has.departed.from.its.natural.state,' the full meaning range of *kangir-* is "meaning, principle, source, headwaters of river, beginning" and "strait of water"). Similarly, Rotuman colexifies 'bay' with 'hollow, cavity, recess.' Yir Yoront *larrngarrp* (containing *larr* "place, site, tract, estate, country") and Khalkha colexify 'bay' with 'bend of river.' Upper Chehalis *s-qíwt=čí* is analyzable as 'CONTINUATIVE-stink=water/river,' and also denotes 'saltwater' and 'sea, ocean,' colexification with 'sea' is also found in Nez Perce.

Other associations include: Yoruba has the (suspiciously long) terms *i-ya-wọ omi òkun sinu ilẹ̀* 'NMLZ-to.branch-to.enter water sea into land' (compare also Japanese *irie* containing *ir-* 'to enter') and *àláfó tí o wà lārin ọ̀wọ̀ mejì* 'space REL 3SG exist between pillar two.' Muna *kolowa* also denotes a 'path in a cave,' and Nunggubuyu *-dun-* also a 'burrow,' while Chukchi (*ka*) *ańee iryin* contains *iryin* 'shore.' Kolyma Yukaghir *öge:dańil* "semi-circular bay without a estuary" contains *ańil* 'estuary.' Welsh *llawryf* also means 'laurel,' and *cilfach* also can refer to a 'nook' or 'recess.' Nuuchahnulth *hita-čus* is analyzable as 'LOC-dig.' Fijian *toba* contains *tō* 'filled, saturated with water or milk' (*toba* also denotes a plant), and for Great Andamanese *tōngmūgu* 'bay,' compare *tōng* 'leaf' and *mūgu* 'face.' There is a semianalyzable term where the identifiable constituent means 'water, fresh water' in Lenakel. Finally, Rotuman *popo* also means 'rotten, decayed' inter alia.

### 5. The Beak

Representation: 73.6%

Motivated: 66.8%

Thereof Analyzable: 15.9%      Thereof Colexifying: 52.4%

Thereof by Contiguity: 3.2%      Thereof by Similarity: 57.5%

Recurrent associated meanings: nose, mouth, snout/muzzle/nozzle, end/point, bird, tooth, lip, prow of canoe, thorn, peck, toe of shoe

To convey the meaning 'beak,' languages overwhelmingly choose body-part metaphors, at least where terms are synchronically motivated. Frequent are associations with the 'mouth.' The association is realized formally mostly by colexification, as found in Buli, Efik, Khoekhoe, Swahili, Baruya, Kyaka (where the relevant term in fact means 'wide beak' specifically), Toaripi, Abzakh Adyghe (among other meanings), Nivkh, Kashaya (the relevant term *haʔbo* is analyzable as /ʔaha-ʔbo/ 'mouth-enlarge/swell.up' and means 'protrusion of the mouth' and 'external mouth'), Yaqui, Hupda (by the term *nəg'öd*, perhaps analyzable as *nəg-öd* 'mouth.related-inside'), Fijian, Malagasy, Rotuman, Samoan, and Tetun

(colexifying also ‘bank, shore;’ secondary associations in some languages that are likely genuinely linked with ‘mouth’ rather than ‘beak’ are not discussed here unless relevant, see section 124 for these). Six languages have morphologically complex terms with ‘bird’ acting as a contiguity anchor, such as Guaraní *gyra-juru* ‘bird-mouth.’ These are Kanuri, Mbum, Yoruba, San Mateo del Mar Huave (where ‘animal mouth’ more specifically is the meaning of the constituent), Guaraní, and Lenakel. Other complex terms where the second element is not ‘bird’ include Ngambay *ngàw-tà* ‘claw-mouth,’ Wappo *naphúce*, analyzable as /nán-phúce/ ‘mouth-poker,’ and Kiliwa *ha?=phaqy*, which contains *ha?* ‘mouth’ and *phaqy* ‘pointed’ (moreover, Hani *meiqtuv* might be analyzable as *meiq-tuv* ‘mouth-peck,’ although *meiq* only occurs with the meaning ‘mouth’ in conjunction with *baq*, a classifier for hollow things, and both *meiq* and *tuv* have other meanings; a semianalyzable term with ‘mouth’ is found in Japanese). ‘Beak’ and ‘end’ or ‘point’ are colexified in Buli, Efik, Ngaanyatjarra (here the term also means ‘face’ and ‘nipple’ inter alia), Abzakh Adyghe (also colexifying ‘beginning’ and other things), Basque (also colexifying ‘front, façade’ and ‘bit smidgeon, a little bit’), Ket (also colexifying ‘top’ and ‘protuberance’), Itzaj, Quileute, Bororo, Guaraní (here also colexifying ‘smoke’ and ‘vapor’), Piro (also colexifying ‘arrowhead’ specifically), Tehuelche (the relevant term is *?or ~ ?ol ~ or*; note the similarity to *?or ~ ?o:r ~ or ~ o:r* ‘tooth;’ there is another phonologically similar term meaning ‘perhaps’), and Yanomámi. There is a semianalyzable term for ‘straight beak’ with the identifiable constituent meaning ‘point’ in Great Andamanese (also, in Cubeo, the root *cāpi-* occurring in *cāpibo* ‘beak,’ when suffixed with a different classifier, assumes the meaning ‘point, headland,’ which is an association found by colexification also in Fijian). At least for some of these languages, the association might be secondary, since the relevant terms in these languages also colexify ‘nose’ in addition to ‘end,’ and this association is in fact the second most frequent lexico-semantic tie to be observed, occurring by colexification in as many as 29 sampled languages (Yoruba, Gurindji, Ngaanyatjarra, Kosarek Yale, Abzakh Adyghe, also colexifying ‘beginning,’ Chukchi, Khalkha, Kildin Saami, Sora, Chickasaw, where the relevant term may also refer to the ‘bill of a hat,’ Itzaj, Lake Miwok, Nez Perce, Quileute (where in fact ‘nose’ and ‘mouth’ are colexified), Wintu (which also colexifies ‘stone, rock’ and ‘bare protuberance generally’), Copainalá Zoque, Aguaruna, where the relevant term also colexifies ‘wax,’ Cavineña, Guaraní, Huambisa, Macaguán, Miskito, Piro, Tehuelche, Tsafiki, Wayampi, Fijian, which also colexifies ‘cape’ and ‘mountain peak’ inter alia, and Hawaiian. An interesting situation is found in Yir Yoront, where “bottom of bird’s beak” is colexified with ‘mouth,’ and “top of bird’s beak” with ‘nose.’ Five languages, Laz, Cheyenne, Kiowa, Santiago Mexquititlan Otomí, and perhaps Bora have analyzable terms involving a constituent meaning ‘nose,’ the contiguity anchor being ‘bird’ in Laz, Cheyenne, and Kiowa, and ‘bone’ in Santiago Mexquititlan Otomí, while in Bora, the relevant term is perhaps analyzable as *túju-wa* ‘nose-CL.PLANK.’ Another frequent metaphor-driven pattern of colexification is that of ‘beak’ with ‘tooth,’ found in Baruya, Kyaka (here, the term denotes a ‘narrow beak’ specifically and also colexifies ‘food’ and “sharp, biting, erosive”), Itzaj (by a different term than that mentioned above; this one also colexifies ‘grain of corn’), Arabela, Bora, Cashinahua, Lengua, and Rama. Compare also Embera *kidhá* ‘tooth’ with *kidá* ‘beak.’ White Hmong has *kaus ncauj* ‘beak mouth’ for both ‘beak’ and ‘fang, tusk, canine



teeth,' while Highland Chontal has *lihay gaga* /lahay ɬaga/ 'tooth bird' and Chayahuita *inaira natë* 'bird tooth.' This pattern is obviously very common in languages of South America. An association exclusively occurring by colexification in the sample is that with 'lip,' found in Swahili, Kyaka (colexifying "wide beak" more specifically), Carrier, Central Yup'ik, Jarawara (where the relevant term may refer to 'fruit' and 'lump' inter alia, likely due to the frequent homonymy in the language mentioned by Dixon 2004), Yanomámi (here the colexification is with 'lower lip' specifically), and Takia (here it is with 'pouting lips' more specifically).

A subsidiary pattern is the association with 'snout,' 'muzzle,' or 'nozzle,' found by colexification in Buli, Khoekhoe, Yoruba, Muna, Khalkha (where *xosiɣu(n)* also may refer to 'chatter,' the 'bow of a vessel,' the 'spur of a mountain,' a 'cape, promontory' or 'peak,' and other things), Haida (also colexifying 'mouth of vessel'), Kashaya, Lesser Antillean Creole French, Nez Perce, Quileute, Copainalá Zoque, Guaraní, Huambisa, Jarawara, and Hawaiian (note that in Nez Perce, Quileute, and Hawaiian, 'nose' is also colexified and in Jarawara also 'upper lip' is; in the other languages the association is not at the same time with other pertinent meanings). In San Mateo del Mar Huave, this association is realized formally by a complex term, namely *ombeay quiec* 'animal.mouth bird.' Possible other body-part metaphors are found in Nunggubuyu, where *dhamunun*<sup>9</sup> is derived from *lhamunun*<sup>9</sup> 'chin' (which is in turn related to *lha* 'mouth'), and the Sko term *lóeto* resembles *loe* 'ear.'

Metaphorical associations not based on the transfer of a body-part term to the 'beak' are the colexification with 'thorn' in Kyaka (where it is colexified with 'narrow beak' more specifically), Bororo, and Lengua, and that with 'prow (of canoe)' in some South American languages, namely Aguaruna, Bora, Guaraní, Huambisa, Piro, and Wayampi (here by the analyzable term *sî-ŋɛ* 'nose/beak-old'), as well as Hawaiian. The cluster in South America is suggestive of an areal pattern of South America, and indeed it may be so, but Blust (2009: 314) reports that this association is also common in Austronesian languages, as suggested by its presence in Hawaiian in the present sample. San Lucas Quiaviní Zapotec and Hawaiian colexify 'beak' with 'handle of pitcher' and 'beaker of pitcher' respectively, and in Khalkha and Hawaiian, there is colexification with 'toe of shoe' (the relevant Hawaiian term may also refer to the "thick pearl of pearl-shell shank").

The only recurrent pattern in the sample that is not metaphor-based is the association with verbs meaning 'to peck,' and even this pattern is highly restricted areally, namely to Western North America. In Upper Chehalis, the 'beak' is called *ɬákʷmri* /ɬákʷ=mn/ 'peck=INSTR,' in Nuuchahnulth *ɬupkyak* is analyzable as /ɬupk-ɣakʷ/ 'peck-instrument,' and in Ineseño Chumash, the same root is used for 'beak' and 'to peck' (but note the evidence from Hani discussed above and that Tetun has the redundant complex term *ibun-tutu-n* 'mouth/beak-peck-SINGULATIVE').

Still other associations include: Khoekhoe *ammi* also means 'deep furrow,' Kyaka *nenge* also 'sharp' inter alia, Greek *rámfos* also 'burner, jet (of light),' Welsh *pig* also 'spout,' and Wintu *só* also 'stone, rock' and *sono* also 'point of rock.' Arabela colexifies 'beak' and 'claw of crab,' and Bororo has one term colexifying 'beak' with 'plait' and 'fibre,' and another one colexifying it with 'sting,' both times among other colexified meanings. Tsafiki uses the same term to refer to the beak of a bird as well as to a 'box,' and Wayampi *sî* also

means 'white.' In Hawaiian, the semantic range of the relevant term *ihu* also includes "mouth or entrance, as of a harbor, river, or mountain pass or gap" *inter alia*.

## 6. *The Bird*

Representation: 98%

Motivated: 20.9%

Thereof Analyzable: 10.9%

Thereof Colexifying: 10.7%

Thereof by Contiguity: 9.2%

Thereof by Similarity: 1.6%

Recurrent associated meanings: animal, chicken/fowl/duck/goose, fly, wing, airplane, penis, insect, feather, game, small/little, sky

Terms for 'bird,' if they are motivated, are often related to the next higher level in the ethnozoological taxonomy, namely 'animal.' In fact, Upper Chehalis, Highland Chontal, Huambisa, Jarawara, as well as Fijian, Rotuman, Kapingamarangi and Samoan, have terms which colexify the two ranks in one term, thus putatively being on the stage of lexical expansion in which terms on the life-form level come to include the unique beginner rank (Samoan also colexifies 'cattle' specifically as well as 'thing' generally; Highland Chontal also colexifies 'bug' and may be an instance of the "wug"-category as defined by Brown 1984; note also that Nunggubuyu, very similarly, has a term also capable of referring to 'terrestrial game animals'). More frequently, however, are complex terms for 'bird' where one of the constituents is 'animal' and the other serves to specify and narrow down the semantic range. Terms where the second element is 'to fly' are found in Buli (*jayirim* /*jaab-yirim*/ 'animal/thing-flying'), Nivkh (*pyi-ŋa* 'fly-animal'), and San Lucas Quiaviní Zapotec (*ma'anyarzu* /*ma'an-rzuhuh*/ 'animal-flies'; the language also has the alternative term *ma'any zhaybàaa* 'animal sky'). However, there are also terms of the derived type, as in Sahu *sosoloro* (*soloro* 'fly'), Nez Perce *weyixniket* /*weye-lixnik-e't*/ 'fly-move.around-AGT,' Nuuchahnulth *maamaati* /*maa-mat-ĩp*/ 'RED-fly-THING....ED,' and Bora *wahpée-be* 'fly-CL.M.S.' Furthermore, semianalyzable terms where one of the constituents means 'fly' are found in Basque, Central Yup'ik, Cavineña, and Toba (here, the identifiable constituent means more specifically 'to fly alone'), and in Wintu, the identifiable constituent of an idiolectal word for 'bird' is "to glide, spread the feathers to glide" (see also Buck 1949: 183 on this association in Greek and some Celtic and Germanic languages). Frequently, terms involving 'wing' are also found (as is also the case in Greek and Sanskrit, see Buck 1949: 183 and § 5.4.2.7.2.). Complex terms with 'animal' being the second constituent are Yir Yoront *minh-puth lon* 'animal-wing with' and Ket *kejassel*, analyzable as /*kēŋ-assel*/ 'wings-wild.animal.' Other terms in which the meaning 'wing' is used include Buin *perekupa*, containing *kupa* 'wing' and presumably *pere* 'to turn, roll' (the term also means 'go hunting for birds' as a verb), Khalkha *zigyr-ten* 'wing-COLL,' and Central Yup'ik *yaqulek* /*yaquq-lek*/ 'wing-one.having' (dialectally meaning 'angel'). There is a semianalyzable term in Miskito, and the association is diachronically detachable also in Dongolese Nubian. In Hupda, the 'bird' is called *hũ-tǣ'h*, perhaps analyzable as 'animal-small,' and a term with the same structure is encountered also in Miskito. Complex terms of the derived type in which the meaning

'feathers' figures are found in Carrier, Kiliwa, and Abipón; note also Cahuilla *wíwikmal* ~ *wikitmal* 'bird' and *wikil* 'feather' (as well as *málmal* 'butterfly?').

A relatively frequent association is with 'chicken,' 'fowl,' 'duck,' or 'goose' occurring by colexification in Basque, Chukchi, Greek (see also Buck 1949: 183 for evidence from Ancient Greek), Khalkha, Welsh, Nez Perce, Central Yup'ik, Abipón (also with 'stork' specifically), Lenakel and Tetun. In addition, in Muna *manu* is 'chicken' (compare the apparent cognacy with the 'animal'-term in Rotuman and other Austronesian languages mentioned above) and *manumanu* is 'bird,' while Sahu *namo diwang* is analyzable as /*namo diwanga*/ 'chicken sky.' This is interesting in the light of Berlin's (1972) report that in Tzeltal the indigenous term for 'bird,' *mut*, experienced a semantic narrowing induced by the introduction of the domesticated chicken by the Spaniards to the effect that its meaning potential was reduced to 'chicken.' In the modern language, 'bird' is expressed by the compound *te?tikil mut* 'forest chicken.' A similar "markedness reversal" (Witkowski and Brown 1983) may have taken place in Sahu and likely also in Muna (but note that there are languages, such as Buli, in which 'fowl' is specifically excluded from the semantic range of the general 'bird'-term).<sup>3</sup> In three sampled languages, Ancash and Imbabura Quechua as well as Tetun, terms for 'bird' are metaphorically extended to 'penis' (a pattern known also in other languages, compare e.g. English *cock*), and in another three, Nunggubuyu, Jarawara, and Lesser Antillean Creole French, terms also mean 'game' ('terrestrial game animals' more precisely in Nunggubuyu). Swahili, Mali, Yir Yoront, and Toba colexify 'bird' with 'airplane' (Mali 'large bird' specifically; on complex terms for 'airplane' on the basis of 'bird,' see section 72). A Khalkha term colexifies 'bird' with 'child,' and the Bwe Karen term *cubápho* contains *-pho* 'child.'

It should also be pointed out that in some languages, 'small bird' and 'large bird' are lexically distinguished; this pattern is found in the sample in Rendille (where the term for 'large bird' also denotes the 'vulture' specifically; compare Sasse 2002: 1053 on this in Cushitic languages in general), Carib, and Yanomámi. In Buli, *jayirim* may also refer to an 'insect,' and the same is true of the Miskito term *daiwan lupia* 'animal small' and the Rotuman term *manmanu ferfere* 'bird/animal restless.' Hani colexifies 'bird' with 'meat,' for which compare the likely relationship between Yei *yarmakər* 'bird' and *makər* 'flesh.'

Further unique associations are rare: In Buli, *nuim* is also the name of a mark made in the skin believed to act as a remedy against a disease caused by a particular bird. Efik *i'nuën* can also refer to a kind of magical conjuration. In Khoekhoe, *anib* ~ *anis* 'bird, cock' is derived from *ani*- 'decorate, adorn, color in, dress up,' Ngambay *yèel* also means 'to feel sick,' while in Kwoma, the relevant term also includes 'flying foxes' in its denotational range and that in Kyaka also 'bat' alongside other meanings (but not flying insects). The relevant Muna term also denotes a children's game involving flying objects, and that in Toaripi also a kind of flying fish. Sko *táng* also means 'fishing net,' while Basque *txori* also means 'thief,' 'mumps,' 'hinge,' 'eyelet,' 'dickey' and 'bun.' Itzaj *ch'iich* also means "pick out (small things), delouse," and for Kiowa *kuatou* compare *tou* 'house.' Wintu *čil* also means

<sup>3</sup> The precise denotatum of English *chicken* is subject to some dialectal variation and there thus may be ambiguities here due to the use of English as the metalanguage.

‘eclipse’ and ‘bear.’ The quite obviously onomatopoetic Wintu term *tili’tilit* also denotes the ‘snipe’ specifically, while Yuki *čim’it* also the ‘blackbird’ specifically. Arabela *shiyojua* contains the classifier *-jua* which is glossed as ‘tierra, volador’ in the source, and for Maxakalí *putuxnāg* compare *putux* ‘heavy.’ Sáliba *llide* also denotes the ‘moriche palm,’ while Wayampi *wila* also denotes a round dance consecrated to birds, as well as, presumably by homonymy, ‘tree, wood.’ Yanomámi *yōō pesi* ‘big bird’ contains *pesi* ‘sheath, nest’ (the term also has other meanings), Bislama *pijin* is rarely used for Bislama, but commonly employed to refer to “Melanesian Pidgin as spoken in Solomon Islands and Papua New Guinea,” i.e. probably Tok Pisin. Hawaiian *manu* also denotes “any winged creature” as well as ‘wing of a kite,’ Lenakel *menuk* is also the name of a child’s illness said to be caused by birds. Manange *ɪɲima* also means ‘gill, ear,’ Sedang *chēm* is also the Sedang name of the star Antares, and Yay has phonologically identical terms for ‘bird’ and ‘to hoe.’

### 7. The Blossom

Representation: 93%

Motivated: 23.4%

Thereof Analyzable 7.7%

Thereof Colexifying: 15.7%

Thereof by Contiguity: 4.2%

Thereof by Similarity: 7.3%

Recurrent associated meanings: plant, bud, fruit, eye, pretty, shine, burst, pod, child, lid/cover, sprout, rose, picture, grass, head, yellow

Lexico-semantic associations of the meaning ‘blossom’ (‘bloom,’ ‘flower’) are manifold, though none of them is particularly frequent cross-linguistically. Six languages, Efik, Khoekhoe, Yoruba, Baruya, Itzaj, and Pawnee colexify ‘bud’ by temporal or provenience contiguity (and Pawnee in addition ‘flower bud’ specifically). Another Itzaj term colexifies ‘flower’ and ‘sprout, seedpod,’ and ‘pod’ and ‘flower’ are also colexified in Rao. Efik (by a term derived from a verb meaning ‘to peel’ and colexifying ‘seed’ inter alia), Kaluli, Kyaka, Sahu, and Samoan colexify ‘flower’ with ‘fruit’ (“inedible tree fruits” in Kaluli specifically; Samoan also colexifies ‘egg’ and ‘products,’ compare the colexification with ‘roe of fish’ in Rotuman). Metaphor-driven conceptualizations, however, are more frequent than contiguity-based ones. Badaga and Hawaiian colexify ‘flower’ and ‘child(ren)’ (the Badaga term also means ‘shoot’ and ‘harvest,’ and the semantic range of the relevant Hawaiian term also includes “tassel and stem of sugar cane” and ‘arrow, dart’ inter alia), Buli and Bororo colexify ‘flower’ and ‘stomach’ (the Bororo term also means ‘liquid, blood’ and ‘shininess’ inter alia, and the Buli term also includes ‘pregnancy’ alongside other meanings in its semantic range, see discussion of connections between ‘stomach’ and ‘womb’ in section 148). In Pawnee, the generic term for ‘flower,’ *kiriktarahkataaru*?, which also, however, denotes the ‘sunflower’ more specifically, contains *kirik-* ‘eye’ and *rahkataar* ‘be yellow’ (‘flower’ is colexified with ‘yellow’ in Malagasy). Furthermore, Wintu *tuh* ‘flower, bloom, blossom’ betrays a lexical connection to *tuh* ‘eye.’ The lexicographer (Pitkin 1984) remarks: “parallel to English daisy, i.e., day’s eye.” Similarly, in Central Yup’ik, the base *uite-* in means ‘to open one’s eyes, to wake up, to bloom (of flowers),’ and parallelly, Burarra

colexifies 'to have eyes open' with 'be in bloom,' by the analyzable term *ga-na ~ ja-na* 'take-see/look.at.'

In Yir Yoront, *par-thaw* is analyzable as 'head/top-mouth,' and in Cheyenne, *tsépêhévêtséavó'o'e* also contains *htséá* 'head;' lexicographers state that a literal morpheme-by-morpheme translation of the entire term would be "that.which.is-pretty-headed-growth" (cf. Central Yup'ik *nau-cetaaq* 'grow-something.used.to.cause.to'), and also note that "[it] is quite likely that there was no historical word for 'flower'" (for 'pretty' as the meaning of a constituent, compare colexification of 'pretty' and 'flower' in Miskito, Ket *aqta daan* is 'pretty/good grass,' and for the association with 'mouth' compare Blackfoot *waqistsisskitsii* 'flower' with *waqitstsi* 'empty one's own mouth of'). The comparison is likely based in both cases on the fact that both 'head' and 'flower' are roundish structures that are found on top of the human body or a twig of a tree or bush or flower respectively. This is further corroborated by Rotokas *kokookua* 'open flower,' which contains *kokoo* "food on plate, something which is placed on top of something else" (see Buck 1949: 525 for a similar association with 'point' and 'edge' in Indo-European). Rama and Yay colexify 'flower' with 'lid' or more generally 'cover' (the Yay term colexifies in addition 'wall of a house'). In Dongolese Nubian, *ǝ́* also means 'fire, flame' and 'light,' and perhaps similarly, Yoruba *itanna ewéko* is analyzable as 'lamp.lightning plant,' Muna *kambea* is derived from *mbea* "glow, shimmer, shine, twinkle," and in Lake Miwok, 'to bloom' is colexified with 'to shine' as of the sun inter alia (see Buck 1949: 526 for Indo-European). Furthermore, for reasons unknown, two languages of Eurasia, Khalkha and Kolyma Yukaghir, colexify 'flower' with 'picture;' in Khalkha, the relevant term also denotes 'smallpox,' and another Khalkha term denotes in addition the concepts 'comb of cock' as well as 'club in cards,' in Kolyma Yukaghir the relevant terms also means 'embroidery,' 'ornament' and 'color.'

Itzaj and Lake Miwok colexify 'flower' with 'sprout.' Comanche *mubutsaakat̚* contains a verb referring to something blowing up or bursting, and 'to bloom, blossom' and 'to burst' are colexified directly in Nunggubuyu and Itzaj. The Dongolese Nubian and Abipón terms also denote the 'rose' specifically (according to Wehr 1976: 1061 this is a pattern of autohyponymy also found in Standard Arabic, so that areal spread or even direct borrowing into Dongolese Nubian is at least a possibility).

Before listing patterns found only in one sampled language each, it must be pointed out (as already done in chapter 3) that a problem with the concept 'flower' is that the English word *flower* is ambiguous between the reading 'blossom' and 'smallish blossoming plant' (German *blüte* vs. *blume*). This hampers systematic analysis, as lexicographers often do not make clear which of the two meanings of English *flower* is meant. It cannot be excluded that colexification with 'plant,' found in Meyah, Badaga, Kildin Saami (here also with 'grass' and 'vegetable'), San Mateo del Mar Huave, Santiago Mexquititlan Otomí, and Central Yup'ik is in some cases due to lexicographers selecting the *blume*-sense of 'flower,' and the same is probably true for terms such as Ket *aqta daan* 'pretty/good grass.' Another point worth mentioning is that terms for 'flower' are in many languages deverbal in nature, with a verb meaning 'to bloom, blossom' being derivationally basic (see again Buck 1949: 526 for Indo-European).

Other associations include: Hausa *huda*, a term of the Kuda dialect for the flower of any plant other than that of the tobacco plant, also means “be slightly split,” “ooze with water,” and ‘to pierce.’ Ngambay *pútu* also means ‘white hair.’ Swahili *ua* also means ‘fence’ and ‘yard, court,’ while Yoruba *òdòdó*, as an adjective, denotes the color ‘scarlet.’ Kwoma *she* also means ‘ash’ and ‘faeces,’ and Kyaka *jingi* also ‘waterfall, rapids’ (for the meaning ‘flower,’ *isa* ‘tree, bush’ may be used in apposition). Nunggubuyu *-dhabag* is derived from *lhabag* ‘feather;’ on the basis of this term, there is the complex verb *-lhababida-* meaning ‘to be in bloom,’ presumably originally containing *-wida-* ‘to be new moon’ (the lexicographer notes that “the semantics make a synchronic segmentation questionable”). Similarly, Buin *paru* ‘feather, plumage’ is also glossed as “(used for) flower.” Sko *pang* in *ripang* ‘flower’ (where *ri* means ‘tree, wood, scales of fish’) is identical segmentally to *pang* ‘pus.’ Cheyenne *véhpôtse* “primarily means ‘leaf’ but is sometimes extended to mean ‘flower’” (compare also Nivkh *en̄v tjomr* ‘flowering leaf’). Muna *wuna*, a term for ‘bloom’ which is however not used for the flowers of most trees and flowers, also denotes “[t]he island of Muna,” as “according to legend it is called Muna (=flower) because the first inhabitants saw a stone with flowers growing on it.” Badaga *hu: ~ u:* also means ‘tail-side of coin,’ and Japanese *hana* also ‘nose, nasal mucus’ (the terms are prosodically different though, and in the ‘flower’-sense, *hana* may be an early borrowing from Austronesian). The literal meaning of *héováéstse tséhetó’o’ee’éstse*, a Cheyenne term for ‘flower,’ is stated to be “various kinds how they grow.” Chickasaw *pakali* is derived from *pakali-* ‘to open, bloom,’ and similarly, Lake Miwok *táke ~ take*, a verb meaning ‘to bloom,’ also may mean ‘to come out, emerge’ inter alia. Itzaj *wach’äl* is derived from *wach’-* ‘to loosen, disentangle,’ and Wintu *lul* is connected lexically to *lul* ‘long cylinder.’ San Lucas Quiaviní Zapotec *gyiia* can also refer to a mixture of corn meal floating on top of tejate (a traditional Oaxaca beverage made from maize and cacao). Aguaruna *yagkúg* is also a man’s name (compare also the segmental similarity to *yagkú* ‘yellow’), Cashinahua *jua* also denotes ‘corn silk,’ and Cayapa *llullu* appears to be a reduplication of *llu* ‘penis, wood.’ Chayahuita *nitërinso* is derived from *nitërin* ‘bear fruit,’ and Jarawara *mowe* also denotes the ‘Brazil nut.’ Tsafiki *lulí* may also refer to a ‘cockscomb’ and is related by unknown means to *luban* ‘red,’ while Toba *lasoviata* is a nominalization of *asovi*, colexifying ‘to bloom’ with ‘to fray, frazzle.’ Yanomámi *horehore* appears to be reduplicated from *hore* ‘hidden below.’ Hani *yeiv* also means ‘not firm,’ Tetun *funan* can also assume the meaning ‘mildew’ and ‘product, interest,’ and *aifunan*, a compound with *ai* ‘tree, plant,’ can also refer to one’s ‘girlfriend.’ Vietnamese *bông* also denotes ‘cotton,’ Yay colexifies ‘flower’ with ‘bone,’ Bislama *flaoa* also denotes “any plant grown for decoration (as against plant grown for food)” and Lesser Antillean Creole French *fle* is also used with the meanings ‘best’ and ‘choice.’

#### 8. The Branch

Representation: 95%

Motivated: 40%

Thereof Analyzable: 12.5%

Thereof Colexifying: 27.5%

Thereof by Contiguity: 9.9%

Thereof by Similarity: 28.3%

Recurrent associated meanings: arm, hand, leg, wing, tributary, horn, knot in tree, tree, fork/divide/separate/branch off, bush, shoot/sprout, point/end, fin, wing of army, descendants, prong, leaf, member, root, bone

The most frequent lexico-semantic association for ‘branch’ is metaphorical in nature and consists of lexical ties with either ‘arm,’ ‘hand,’ or both (assuming in the following discussion that terms glossed as ‘limb’ do not exhibit any of the patterns of colexification). This pattern is also attested diachronically in the history of Germanic according to Buck (1949: 523). Formally, both colexification and morphologically analyzable terms are amply attested. All in all, the association with ‘arm’ is found in 29 sampled languages, by colexification in Koyraboro Senni, Buin, Burarra, Kwoma, Mali, Ngaanyatjarra, Khalkha (more generally, ‘limb’ is colexified here), Ineseño Chumash, Itzaj, Kashaya, Kiliwa (by the analyzable term *t+haq=tay* ‘OBJ+bone=be.large;’ ‘branch’ and ‘bone’ are colexified, inter alia, in Bororo), Quileute, San Lucas Quiaviní Zapotec (colexifying ‘upper arm’ and ‘upper arm of front leg of animals’ specifically), Carib, Cashinahua, Jarawara, Bislama (among other colexified meanings), Great Andamanese, Fijian (also colexifying ‘upper arm’ specifically), Hani (also colexifying ‘branch of clan’), and Lenakel, and by analyzable terms in which ‘tree’ acts as a contiguity anchor, as in Yir Yoront *yo-puth* ‘tree-arm,’ in Anggor, Dadibi, Yir Yoront, Highland Chontal, Santiago Mexquititlan Otomí, Chayahuita, Kaingang, and Yanomámi; furthermore, a semianalyzable term where one of the constituents is ‘arm’ is found in Yana. With ‘hand,’ colexification is found in Koyraboro Senni, Yoruba, Buin, Kwoma (also with ‘handprint’ and “artist’s personal style of painting and carving”), Mali, Ineseño Chumash, Itzaj, Quileute, San Lucas Quiaviní Zapotec, Jarawara, Bislama (again, among other colexified meanings), and Lenakel, and analyzable terms are present in Mbum (*ndòk-kpù* ‘hand-tree’), Anggor, and Kosarek Yale. Moreover, there is a semianalyzable term in Pipil. Now, languages frequently colexify ‘hand’ and ‘arm’ (Brown 2005b), and where the above groups overlap, this is due to such colexification.

Other complex terms with ‘tree’ as contiguity anchor include One *aila plana* ‘tree two,’ San Mateo del Mar Huave *omal xiül* ‘point tree’ (for which compare colexification of ‘branch’ with ‘pointed, pointed object’ in Abzakh Adyghe and with ‘pointed, sharp’ inter alia in Muna), and Tetun *ai-sorun* ‘tree-side.’ Furthermore, Nuuchahnulth *-maptmapt* is reduplicated from *-mapt* ‘plant, bush, tree,’ and semianalyzable terms in which one constituent is identifiable as ‘tree’ (and sometimes ‘wood,’ see section 65.) are found in Berik, Sko, Waris, Rama, and Bwe Karen. However, ‘tree’ cross-linguistically not only serves as a contiguity anchor in complex terms, but colexification of ‘tree’ and ‘branch’ by spatial contiguity, or more precisely meronymy, is also attested, namely in Ngambay (where many other tree-related meanings are colexified as well, among them ‘forest’), Buin (colexifying ‘small tree’ and a particular tree species more precisely), Welsh, Blackfoot, and Kiliwa. Similarly, in Efik, *okpüt* also colexifies “branches of the young trees lopped off in clearing a plantation” as well as “the trees themselves being left as yam sticks.” Moreover, Dongolese Nubian, Muna, and Khalkha colexify ‘branch’ with ‘shoot, sprout,’ and Haida *hlq’a7ii* also means ‘bush,’ “stem (raceme or panicle) of berries or of blossoms grow-

ing on bushes” and ‘supporting frame’ inter alia; the association with ‘bush’ is also found in Middle Eastern and Southeastern Tasmanian and Nuuchahnulth.

Interestingly, an alternative to ‘hand’ and/or ‘arm’ is colexification with ‘leg,’ as found in Buli, Khalkha (‘limb’ generally is colexified here), Arabela, and White Hmong (in Buli also with ‘hind leg,’ ‘foot’ and other meanings specifically, and in Arabela also with ‘handle of pot,’ ‘backrest of hammock,’ ‘grip of bag,’ and other things), and with ‘foreleg’ specifically in Buin, Ngaanyatjarra, San Lucas Quiaviní Zapotec, and Fijian (in Buin also with ‘pectoral fin,’ ‘fin’ is colexified with ‘branch’ inter alia also in Hawaiian). As in Buli, Nunggubuyu colexifies ‘branch’ with ‘hind leg, upper leg’ specifically, and has the redundant complex term *-n<sup>u</sup>u-dhalbar* ‘-guts-upper.leg/hind.leg,’ for which compare Maxakalí *mīm-māg* ‘wood-small.intestine.’ Similarly, colexification of ‘branch’ and ‘leg of insect’ is found in Kwoma, while Wappo colexifies ‘branch’ with “many-branched, many-legged.”

Other notable metaphor-driven patterns of colexification are that with ‘horn’ or ‘antlers,’ occurring in Hausa, Ngambay (colexifying also other meanings), Basque, Kildin Saami (more specifically, “branching at the antlers”), Abipón (here, ‘jump’ is colexified in addition, and there is a redundant complex term involving ‘wood’), and Toba, that with ‘wing’ in Dongolese Nubian, Buin, Burarra, Kwoma, Ngaanyatjarra, Cashinahua, Jarawara, and Fijian (Dongolese Nubian, notably, is the only of these languages in which the association with ‘wing’ is surely genuine, that is, in which the relevant terms do not also colexify ‘arm’), and that with ‘tributary’ of a river in Khoekhoe, Baruya, Basque, Khalkha, Arabela, Jarawara, and Toba (and in Burarra, *mernda* ~ *marnda* and *gurnjirra* also denote the ‘creek’ itself, an association that is at least etymologically also detectable in Nunggubuyu and is also attested in Germanic according to Buck 1949: 524). Semantic extension to ‘wing of an army’ is documented for Hausa and Hawaiian (and to ‘division’ in Lesser Antillean Creole French), to ‘descendants’ in Hausa and Khoekhoe (in which latter the relevant term also can refer to a ‘tiller,’ ‘sucker’ and ‘water-shoot’ and “leg of skin/hide”), to ‘member,’ as of e.g. a society, in Khalkha and Hawaiian (in both languages alongside other meanings), and to ‘prong’ in Khalkha and Rotuman (in Khalkha also to ‘component’ generally, as well as ‘department,’ ‘detachment’).

Terms for ‘branch’ are also frequently associated lexically with meanings such as ‘to fork,’ ‘to divide,’ ‘to separate,’ or ‘to branch off’ (see Buck 1949: 523). Kaluli *i gasa* is analyzable as ‘tree things.that.separate’ (the language has another term, *i elé*, where *elé* is identical segmentally to the word for the ‘moon’), Kyaka *paka pingi* as ‘branched/divided root/rudiment/base,’ itself denoting a ‘division’ or ‘fork’ as well, and Pawnee *rakitwai’u* as */rak-kita-wi-u’/* ‘tree/wood-be.branching-??-NOM.’ Khalkha has a derived term, and such associations are also found by colexification in Badaga, Kildin Saami, Rotuman, and Sedang (furthermore, Sahu *sasalanga* colexifies ‘fork in tree, branch, fork in road,’ and Muna *tuna* also means ‘fork of fingers and toes’ as well as ‘to bud’). Furthermore, five languages of Western North America, Carrier, Upper Chehalis, Lake Miwok, Nez Perce, as well as the Nunivak Island dialect of Central Yup’ik colexify ‘branch’ and ‘knot in tree’ or ‘knot in wood,’ and two sampled languages, Gurindji and Nuuchahnulth, colexify ‘branch’ and ‘leaf’ (Gurindji more specifically ‘leafy branch’ while Nuuchahnulth also uses the relevant term to denote the ‘Kinnikinnick Berry’). Ngaanyatjarra colexifies ‘side root,’ and Buli ‘root’



generally (compare the association with 'root' in Indo-European mentioned by Buck 1949: 523).

Other patterns found in the data are: Efik colexifies 'branch' with 'root shoots of cereals,' and another Efik term, *ñ'kōk*, is derived from *kōk*, meaning 'to overlay, overlap' and 'to heap up' inter alia. Khoekhoe *llnâub* also denotes the 'brim of a hat,' Rendille colexifies 'branch' with 'notch' as well as '(male) dog,' and Yoruba with 'spray.' Muna *karagha* also denotes a "large natural drain, ravine" inter alia, *tuna* also means 'bud' as well as 'to bud, sprout,' and *ragha* also 'plot, field, piece of land.' Ngaanyatjarra *mirna* may also refer to a "pannikin, mug, cup," 'sleeve of dress,' and "steering arm or wishbone of vehicle." Sahu colexifies 'branch' with "midrib of a palm frond," while Abzakh Adyghe *q<sup>ə</sup>ətame*, a formally redundant compound of *q<sup>ə</sup>ə* 'branch' and *tame* 'wing,' may also refer to 'grape,' and *q<sup>ə</sup>əre* also means 'hard, rough, raspy.' Badaga *mole* also means 'nail' and 'peg' and *sui* also 'whorl of hair.' Basque *adar* also can refer to a 'bedpost' or 'chairpost,' as well as a 'lineage.' Khalkha *nailzayur* ~ *nailzur* is analyzable as /*nailza-yur*/ 'swing-INSTR' (the variant *nailzayur* also means 'glanders') and another Khalkha term colexifies 'branch' with 'tendrils,' while Welsh *cainc* also means 'tune, song' and 'knot.' Nuuchahnulth colexifies 'branch' with 'rind, bark' as well as, by another term, with 'plant.' Wintu *lob* is said to be connected to *lEw* 'oscillate' as well as to meanings like 'hanging, overhang.' The base *ava-* in Central Yup'ik *avayaq* occurs in other terms with the meaning 'son.' Aguaruna *tsajám(pe)* also means 'finger,' Cashinahua *punyan* also 'sleeve,' Embera *həwáte* also 'stick pole' (the meanings are associated with different genders), and Jarawara *mani/mano*, colexifying already 'hand,' 'arm,' 'wing' and 'tributary,' furthermore may refer to a 'watch band' and 'bunch.' Macaguán *pebukrán* might be semianalyzable (*pebú*, 'fruit'), and Sáliba *anojahñojo* appears to contain *anojaha* 'shoulder.' Wayampi colexifies 'branch' with 'head' and other meanings, and Great Andamanese *âkàchâti* appears to be derived by prefixation of a possessive prefix from *châti*, a term for a 'yam species.' Bwe Karen colexifies 'branch' with 'narrowest part of something,' and Fijian inter alia with 'store of the house' and 'page of book.' Kapingamarangi *manga* also means 'to turn brown from heating' and 'to tan,' Hawaiian *lālā* colexifies 'coconut frond' and 'to diverge,' among many other meanings, and *mana* also means 'mana,' 'power,' and 'variant, version,' again among other meanings. Finally, Kapingamarangi colexifies 'branch' with 'branch of government.'

### 9. The Bud

Representation: 51%

Motivated: 53.5%

Thereof Analyzable: 19.7%

Thereof Colexifying: 33.8%

Thereof by Contiguity: 10.5%

Thereof by Similarity: 7.9%

Recurrent associated meanings: shoot/sprout, flower/blossom, to sprout, seed, eye, young leaf, child, tobacco, pimple, button, fruit, foam, egg, burst open

Terms for the 'bud' are sometimes associated lexically with the meaning 'to sprout.' Yoruba has *ēhu ohun ògbìn* 'sprout thing plant,' Berik *fas ferwesini* /*fas ferwesili*/ 'zero/NEG sprout,' in Chayahuita *pichopitërinso* is lexically related to *pichopitërin* 'to sprout,' Tetun

has *ai-tubuk* ‘plant-to.sprout,’ while ‘bud’ and ‘to sprout’ are colexified directly in Hausa and Muna. Colexification with ‘shoot, sprout’ (and sometimes also ‘seedling’) is attested in Kyaka, Ngaanyatjarra (also with “small plants emerging from the soil” as well as “budding grain, unripe grain” specifically), Rao, Badaga, Basque, Welsh, Pipil (the relevant term *-mulinka* is derived from *-mulu:ni* ‘to swell up, dry out’), Wappo, Lesser Antillean Creole French, Aguaruna, Arabela, Bora, Embera, Guaraní, Ancash Quechua, Yanomámi (which also colexifies ‘new branch’), and Tetun, in which latter some of the relevant terms are analyzable: *ai-bo’as* ‘plant-burst’ and *ai-dikin* ‘plant-shoot.’ Ten languages in the sample showcase associations between the meanings ‘bud’ and ‘flower’ or ‘blossom.’ These are Efik, Khoekhoe, Yoruba, Baruya, Itzaj, and Pawnee by colexification (in Khoekhoe and Pawnee, relevant terms are deverbal derivatives from verbs meaning ‘to bloom,’ while in Itzaj, *wach’äl*, which can also mean ‘tassel,’ seems to be derived from *wach’* ‘to loosen, disentangle;’ moreover, the Samoan term may optionally be enhanced by an element meaning ‘to open’). As for complex terms, the Basque term is metaphorical and based on ‘eye’ (*lore-begi* ‘flower-eye,’ note also direct colexification of ‘eye’ and ‘bud’ alongside other meanings in Ancash Quechua and Hawaiian as well as Piro *yhale-xi* ‘eye-fruit/seed/DIM’ which, in addition, colexifies ‘ankle’ and ‘belt buckle’), while Kyaka *jingi ene* is analyzable as ‘flower unfinished,’ Wayampi *poti-ya?i* as ‘flower-child’ and Kapingamarangi *akai modo* as ‘flower unripe.’ Furthermore, three sampled languages, Sko, Jarawara (by the analyzable term *aka-bori* ‘wear-container/nest’), and Hawaiian, colexify ‘bud’ and ‘young leaf,’ and in two sampled languages, Buli and Tetun, a lexical association with the meaning ‘burst’ or ‘burst open’ is found - in Buli by colexification (also with ‘to pierce,’ ‘to rise’ inter alia), and in Tetun by the complex term *ai-bo’as* ‘plant-burst;’ furthermore, Piro *muhip-werekatu* also appears to contain *hipwere* ‘to burst open.’ Colexification with ‘fruit’ is found in Efik (by a term derived from a verb meaning ‘to peel’ inter alia) and Sáliba (also with ‘seed’ in this case, as in Efik and Lesser Antillean Creole French, which latter also colexifies ‘germ’ and ‘rose-bud;’ note also Oneida *yonáhute?*, containing *-näh-* ‘seed, grain, oats’ and *-N-ut-* ‘be attached,’ and the Piro term already mentioned above). Efik and Toaripi colexify ‘bud’ with ‘tobacco’ (“trade tobacco, tobacco twist” specifically in Toaripi), and Khoekhoe and Ancash Quechua with ‘foam’ (the Khoekhoe term denotes the bud of *Acacia Watkins* specifically).

Further metaphorical associations include that with ‘child,’ occurring by colexification in Tuscarora (here also with ‘embryo’) and Hawaiian (where the relevant term also has a range of other meanings) and by a morphologically complex term (*poti-ya?i* ‘flower-child’) in Wayampi (note also colexification of ‘bud’ with ‘descendant’ in Wichí), as well as the colexification of ‘bud’ and ‘pimple’ in Fijian and Lesser Antillean Creole French (also with ‘protuberance’ generally inter alia in Fijian). In two languages of the sample, Lesser Antillean Creole French and San Lucas Quiaviní Zapotec, the relevant terms also mean ‘button;’ this is likely due to copying of the pattern of colexification in the lexifier language French for the former and to that in the donor language Spanish (*btoony* is a loanword from Spanish *botón*) in the latter language. Finally, Kaingang *jumĩ k̃y jẽ* ‘flower bud’ looks as if it contains *jẽ* ‘tooth,’ and the same is true of Meyah *ofoncowú* ‘to bud’ (*ofon* ‘tooth’).

Other associations include: Efik *i'kõñ* has a wide range of meaning, including 'grass, herbs,' 'vegetables of all kinds,' 'leaf,' and 'tobacco.' Hausa *dum'baru* also means 'lips' as well as "first coming-through of a tooth," while *toho* also denotes a particular drum. Noni *com* 'to bud' also means 'to color,' Berik *koksa* also 'to grow,' Kyaka *kene* may also refer to a "small edible crab, found in some streams," while the Muna term *tuna* may also refer to a 'twig' or 'branch' as well as to the 'fork of fingers and toes.' Muna *lasi* also means 'to weld' (this sense is due to borrowing from Indonesian/Dutch). Ngaanyatjarra *yurnturntu* is also used with the more specific meaning "budding grain, unripe grain." Rotokas *koko ruu* is a complex term consisting of the constituents *koko* 'pour out, spill' as well as 'dish out, portion out' and *ruu* 'enclose,' a variant of which is *kokooko ruu*, where *kokoo* means "food on plate, something which is placed on top of something else." Toaripi *kuku* assumes the meaning 'unopened' when occurring in compounds, Sahu *boro* seems to be derived from *woro* 'wide, spacious,' and the primary meaning of Yir Yoront *nhapn* is 'egg' (note also that Great Andamanese *ârmôl* appears to be derived from *môl-o* 'egg,' and that a diachronic association between 'bud' and 'egg' is likely in Wintu). Badaga colexifies 'bud' with 'hair knot, braid,' and Basque *ernamuin* also means 'outbreak.' The meaning of Khalkha's *comurliy* ranges over "bundle, bunch, bouquet" as well as 'anthology,' whereas *nakija* is 'shoot of grass, leaf bud' (compare *naki-* 'to bend; to cave in, be shaky?'). Upper Chehalis *mayspáqññ* contains *máy* "begin to, just now, recently, just a little while ago, hardly" and *páqin* 'to bloom.' Arabela *jiyoto* colexifies 'bud' and 'nut,' and Embera *dodháa* also means 'cocoon.' Miskito *kiama namika uruwan ba* is not fully analyzable, but the element *kiama* can also refer to the 'ear,' while Ancash Quechua colexifies 'bud' with 'yerbasanta,' a type of tree. Toba *l̥̥axataxa* is derived from a verb meaning 'to raise, grow,' Fijian has *kovukovu*, with *kovu* 'to tie up' being the reduplication base. The term is also used with reference to an enclosed portion of land. Hani colexifies 'to bud' with 'to weed with hoe.' The semantic range of Hawaiian *liko* also includes 'shining, sparkling' and 'fat' inter alia, while 'ōpu'u is derived from *pu'u* with the basic meaning 'protuberance.' Samoan *moemoe* is a reduplication of *moe*, meaning 'sleep,' and Sedang *romôe lóang* also may refer to the "soft, tender part of a plant." Finally, Tetun *tubun*, adjectivally, also means "young, fledgling."

#### 10. *The Cave*

Representation: 84%

Motivated: 53.3%

Thereof Analyzable: 24.3%

Thereof Colexifying: 29.2%

Thereof by Contiguity: 37.9%

Thereof by Similarity: 4.5%

Recurrent associated meanings: hole, stone/rock, den/lair, earth/ground, tunnel, valley, pit, opening, house/dwelling, ledge/overhang, in/inside, hill, mouth

As for the 'cave' ('cavern,' 'grotto'), there is one motivational pattern that is by far the most frequent and outnumbers all others in terms of the number of languages it occurs in, namely complex terms of the lexical type where one of the constituents means 'rock' or 'stone' and the other 'hole.' Structurally, these terms obviously vary from language to language. For instance, Hupda has the compound *păç m̥ay* 'stone/hill hole/house,' Yoruba

has the phrasal term *ihò inu àpáta* ‘hole in rock,’ Cubeo *cūracobe* consists of *cūra* ‘stone’ and *-cobe*, a classifier for hole-like objects, and in Oneida, *yotstahlaká'lute?* consists of the verb *-kahlut-* ‘be an opening, be a hole’ which has incorporated the root *-itstahl-* ‘rock, mountain, outcropping’ (there is an alternative term where the incorporate is *-uhwatsy-* ‘land, earth, ground’) and accompanying grammatical material, but the source concepts are remarkably constant cross-linguistically. All in all, this pattern is found in 20 sampled languages all over the world: Yoruba, Berik, Toaripi, Basque, Highland Chontal (with some phonological deviations), Haida, San Mateo del Mar Huave, Oneida, Bororo (where *ia* in *ia-ri* also means ‘mouth’ and ‘opening’ more generally, compare also Baruya *simaanga* /*silamaanga*/ ‘stone-mouth’ and the semantic extension of ‘mouth’ to ‘opening’ in general discussed in section 124), Cubeo, Guaraní, Hupda, Rama, Wayampi, Hani (where the association is also realized by means of a classifier), Lenakel, White Hmong, Tetun, and Bislama. In addition, San Mateo del Mar Huave and Tuscarora have complex terms based on roots meaning ‘hole’ and ‘hill,’ while Cheyenne *tsévé'evótoo'e*, literally, according to the source “that which is a concave hole,” can be extended to mean ‘vagina’ or ‘anus.’ A derived term featuring a lexical constituent meaning ‘to make a hole in the ground’ is featured in Great Andamanese, a semianalyzable term one of the constituents of which is ‘hole’ is found in Rama, and one with ‘stone’ is featured in Kemtuik. Alternative complex terms based on ‘stone, rock’ are Itzaj *'aktun*, perhaps /*'aak-tun*/ ‘turtle-rock’ and Chayahuita *na'pitëana*, containing *na'pi* ‘rock’ and *-ana*, a classifier conveying the notion ‘around’ (there is also the term *cari-ana* ‘precipice-CLASS.AROUND’). In Malagasy, there is a complex term *lavabàto*, consisting of *vàto* ‘stone’ and *làvaka*, which in fact colexifies ‘hole,’ ‘pit,’ and ‘cave,’ and it is not the only language where the association between ‘cave’ and ‘hole’ or ‘hollow’ is formally by colexification: Efik, Sko, Badaga, Khalkha (colexifying also ‘empty’), Kildin Saami, Sora, Acoma (by the analyzable term *pāniizeēša* containing *pāni* ‘bag’ and *zeeša* ‘place where’), Cahuilla, Upper Chehalis (where *ʔackwál[a]xʷɬ* contains the root *kʷálxʷ-* ‘hollow’ and a marker for stative aspect), Cheyenne, Comanche, San Mateo del Mar Huave, Itzaj, Lesser Antillean Creole French, Quileute, Xicotepec de Juárez Totonac, Wappo, Wintu, Aguaruna, Cavineña, Cayapa, Huambisa, Ancash Quechua, Toba, Yanomámi, Fijian, Hawaiian, Kapingamarangi, Mandarin, and Takia all feature terms that are likewise of this type (Toba and Yanomámi also colexify ‘opening’ generally). In analyzable terms, an alternative pattern is for the constituent other than ‘hole’ to mean ‘earth,’ ‘ground’ or ‘land’ rather than ‘stone,’ as in Cayapa *tu-juru* ‘earth-hole.’ Such terms are also found in Kaluli, Chickasaw, Oneida, Pawnee (where the relevant term also colexifies ‘cellar’ or ‘storm cellar’), Tuscarora, Bora, and Guaraní (where an additional element meaning ‘big’ is present), and a semianalyzable term where one of the constituents is ‘earth’ is also found in Huambisa. A semantic association loosely related to the complex revolving around the meanings ‘stone’ and ‘rock’ in some languages is that with ‘house’ or ‘dwelling’: Kyaka has *kana anda* ‘stone/rock house’ (with both constituents also having other meanings), Sahu *ma'di sasa'du* ‘cavern’ is analyzable as ‘stone/rock ceremonial.house’ (more precisely, *sasa'du* is glossed as “ceremonial house with octagonal roof present in each original Sahu village”), and, as already seen, in Hupda *māy* in *păç māy* colexifies ‘hole’ and ‘house.’ The association with ‘house’ or ‘dwelling’ is found by colexification in Kwoma, while Comanche colexifies ‘hole,’

'cave' and 'room.' Burarra (where the relevant term *gun-ngarnama* is analyzable as 'CLASS.DOMESTIC-inner.thigh') colexifies 'room' and 'wall.'

There are also other unrelated lexico-semantic associations for terms meaning 'cave.' Sora *luḡərən* colexifies 'inside,' while Chickasaw *yaaknanonka* is analyzable as /*yaakni*-'anonka'/ 'earth-inside,' Bwe Karen *lo-bú* as 'stone-in' and *ḡ-bú* 'in a cave in the cliffs' as 'cliff-in.'

Eleven sampled languages, namely Hausa (one of the two relevant terms denotes "a hole in the ground where mat-makers sit and work" inter alia), Noni, Ngaanyatjarra, Badaga, Khalkha, Nivkh, Kildin Saami, Cheyenne, Arabella, Toba, and Hawaiian colexify 'cave' with 'den' or 'lair' of an animal, and in five sampled languages, Ngambay, Khalkha, Nivkh, Nez Perce (colexifying also 'bluff' and 'rockshelter') and Pipil, the relevant terms also denote a 'ravine' and/or a 'valley.'

Six sampled languages of the Americas, Haida, Oneida, Quileute, Wintu (where the relevant term *holoq* contains *holo* 'dark'), Aguaruna and Cashinahua colexify 'cave' and 'tunnel,' whereas Khalkha, Sora, Itzaj, Hawaiian, and Malagasy colexify 'cave' with 'pit.' Furthermore, three of the sampled languages, Khoekhoe, Rendille, and Kosarek Yale have terms for 'cave' (in Kosarek Yale more specifically 'cave (where bats dwell)') which can also refer to a 'ledge' or an 'overhang.' The Kosarek Yale term is *bubun*, for which compare *bubu* 'liver, carrying part, point of support.'

Other less common patterns include: Efik *aba* also means 'more, gain' and 'forty, fortieth,' Ngambay *wél* also means 'rapidly,' Dongolese Nubian *ḡār* also 'shore, bank, water-side' and 'to crush (grain)' (in the sense 'cave' it is a loanword from Arabic), Lavukaleve colexifies 'to burst,' and Meyah *mansú* can also refer to a hollow in a tree. Muna *lia* also denotes 'stinging hair on leaves, bristles.' Rotokas *kakiua* contains *kaki* 'opened, cracked, split open' and the classifier *ua* for narrow objects. The same suffix is present in *tariuea*, where the lexical root means 'to encircle.' Sko *long* also means 'key,' and Yir Yoront colexifies 'cave' with 'jail.' Abzakh Adyghe *nəbye* also means 'nest,' Badaga colexifies 'place' generally, and Basque *leize* can also refer to an 'abyss,' 'chasm,' or 'depths' in general. Khalkha *kebtəsi* also means 'uterus' and 'placenta.' Sora *'rupa:n* also means 'pit' and 'pool in stream' and is derived from *'rupa:-* "to form pits and hollows on account of rain." Tuscarora *yunhúčę?* is built around the lexical root *-nhúčę-* meaning 'be a corner' and also means 'cape' and 'nook,' while Wappo *hóča* also means 'sweathouse' or 'dancehouse.' San Lucas Quiaviní Zapotec *cweeb* is also the name of a children's game, and Copainalá Zoque colexifies 'hole in ground, cavity.' Aguaruna *juwaínu* can also mean 'island,' 'opening,' and 'drilled,' and Aymara *putu* also can refer to a small hole in the ground where produce is stored. Cashinahua *bean debu* contains *debu* 'point, end, source of river,' while Cavineña colexifies 'cave' (and 'hole') with 'well.' Huambisa *waa* can also refer to a 'partridge,' Ancash Quechua *machay* also means 'to get drunk,' and the Piro term for 'cave,' *mahka*, can also refer to a 'hill, ascent.' Tehuelche *ma:* is also one of the variants of a verb meaning 'to kill.' Hawaiian *ana* colexifies 'larynx' inter alia, *lua* in the same language also denotes 'grave,' 'mine,' and 'crater' alongside other meanings, and *pao*, again alongside other meanings, also '(arch of a) bridge' as well as 'to scoop out, dub out.' Kapingamarangi *lua* is also the numeral 'two.'

### 11. The Clearing

Representation: 32%

Motivated: 56%

Thereof Analyzable: 42.7%

Thereof Colexifying: 13.3%

Thereof by Contiguity: 45.6%

Thereof by Similarity: 2.7%

Recurrent associated meanings: open/clear/unobstructed, meadow/lawn/field/glade, valley, cut, gap, tree, nothing

Terms for ‘clearing’ frequently make reference by verbal elements to notions such as ‘to open up,’ ‘to clear up,’ or contain other lexical items meaning ‘clear,’ ‘open,’ or ‘unobstructed.’ Such terms are found in eleven sampled languages, namely Efik (where the relevant tem can also refer to ‘openness’ as a character trait), Nunggubuyu, Rotokas, Welsh, Upper Chehalis, Nez Perce, Bororo (where ‘clear’ is colexified with ‘light, bright,’ inter alia; Fijian colexifies these meaning directly with ‘clearing’), Miskito, Yanomámi, as well as in Hawaiian (the term is curious: *pā-pū* also means ‘fort, fortress’ and is in this sense literally translated by lexicographers as ‘gun enclosure.’ It also means “plain, clear piece of ground; clear, unobstructed, visible, in plain sight, directly confronting,” and since both constituents have very many meanings, it is hard to tell which of these are relevant for conceptualization), and Samoan. For instance, in Rotokas, the term for a ‘clearing’ is *koraua*, containing *korau* ‘clear, unobstructed’ and the classifier for narrow objects *ua*, and in Bororo *boe ali* is analyzable as ‘thing clear;’ Nunggubuyu *abaḷa-yarayara* ‘open.area slender’ also denotes the ‘clear sky.’ Itzaj has *jāwänka’il* /*jāwän-kab’il*/ ‘open-world-ABSTR’ for both ‘clearing’ and ‘plaza’ (alongside *petexka* /*petex-kab*/ ‘circle-world’). Moreover, Great Andamanese has *êr-tâlimare* ‘place-clear.’

Four sampled languages, Efik, Kaluli, Kwoma, and Hawaiian, have terms referring to the fact that a clearing may come into being by the cutting down of trees (although ‘tree’ is not always present as a contiguity anchor): Efik *eritem* is derived from *tem* ‘to cut down’ (the term can also refer to a ‘boiling, cooking,’ which is due to *tem* also having the corresponding verbal semantics ‘to boil, cook’ inter alia), Kaluli has *i-kuwo*: ‘tree-cut,’ Kwoma *aka tobo me yii* ‘house trunk tree cut’ and *yii poy* ‘cut completed,’ and Hawaiian *kāhanahana*, which is related to *kahana* ‘cutting.’ In another two languages the terms make reference to the fact that on a clearing there is no plant cover: Baruya has *ku’maaya* /*kutawi-maaya*/ ‘inside-none,’ meaning ‘open area, clear area, area without grass’ and Carrier *hwozzai-ket* ‘nothing-on’ (this term colexifies ‘glade’). In another three languages, terms utilize in some way the fact that a clearing improves visibility: Basque *argi-une* ‘light-space/moment’ (also indeed denoting a “moment of light” or a “moment of ludicity”), Bororo *baru-gwa* ‘sky piece,’ and Fijian *talāsiga*, made up of *tala* ‘to clear away rubbish, transplant, load cargo’ and *siga* ‘day, daylight, sun,’ in addition Great Andamanese term *elôtwālnqa* colexifies ‘clearing in jungle’ with ‘view.’ Seven sampled languages colexify ‘clearing’ with either ‘meadow,’ ‘glade,’ ‘lawn,’ or ‘field.’ These are Badaga, Carrier, Kashaya, Lake Miwok, Pawnee, Tuscarora, and Central Yup’ik. Five languages, Khalkha, Chickasaw, Itzaj, Lake Miwok and Pawnee (here by a term that can be literally translated as “flat ground place,” also meaning ‘flat ground, prairie,’ and ‘low level area,’ inter alia)

have terms that colexify ‘clearing’ and ‘valley,’ ‘glade,’ or a similar meaning, the motivation probably being similar to that for colexification of ‘forest’ and ‘mountain’ (see also sections 26 and 39 as well as § 6.3.). Finally, Badaga and Itzaj colexify ‘clearing’ with ‘gap.’

Other associations include: Khoekhoe *!gorab* ~ *!goras*, also denoting a “bare/denuded area (as when burnt down/mowed),” is derived from *!gora* ‘barren, denuded, bald,’ and Noni *ηkase* consists of *kase* ‘to clear a field’ and the noun class prefix *η-*. Buin *aarono* also can refer to the ‘open sea’ (in old style) or an ‘open space’ in general, while Baruya colexifies ‘clearing’ with ‘bald.’ Ngaanyatjarra *yurirl(pa)* is similarly used to refer to an ‘open place,’ but is also extended to mean ‘room, space’ and ‘outside,’ and Nunggubuyu *abaḷa* can also refer to a ‘ritual clearing.’ Badaga *naṭṭa* ~ *natta* means “common wasteland, common place used for non-agricultural purposes,” and *tera* also “chance, opportunity, course.” Khalkha *cabcijal* seems to contain *cabci-* meaning inter alia ‘to blink’ and ‘to chop, to hew,’ and Sora has terms that colexify the meanings ‘hill,’ ‘clearing on a hill,’ but also ‘forest;’ they are derived from a verb meaning ‘to make a clearing on hill, cultivate on hill.’ Chickasaw *kochchaafokka* is analyzable as /*kochcha*’ aa-fokha-’/ ‘outside LOC-be.in-NMLZ,’ and denotes “an open place out in the middle of nowhere” generally, including a ‘valley.’ Haida colexifies ‘muskeg’ and ‘bog,’ and Kiowa *’H̄t̄ǣt̄’eidl* contains *’H̄* ‘wood’ and *t̄ǣ* ‘be smooth;’ the unknown element *t̄’eidl* is probably related to *t̄’H̄* ‘to cut several’ according to the consulted source. Itzaj *yam* may also refer to a ‘groove’ or ‘canal,’ and yet another Itzaj term, *paak*, also means ‘stack, packet.’ Central Yup’ik *cangurneq* consists of the root *cangur* ‘to lack symmetry’ and the postbase (cf. § 4.4.2) *-neq* ‘area.’ Carr’*ilquq* ~ carr’*ilqaq*, another Central Yup’ik term for ‘clearing,’ contains *carrir-* ‘to clean’ and the postbase *-quq* ‘one that is.’ Bororo *boe giguduru* appears to contain *boe* ‘thing’ and *gigudu* ‘dust.’ Fijian *lalama* also means ‘transparent, too thin (of cloth),’ Hawaiian *paia* also ‘wall, side of a house,’ as well as “walled in” by vegetation, and Samoan *lafo* inter alia also denotes a “plot of land (made ready for planting).”

## 12. The Cloud

Representation: 95%

Motivated: 41.2%

Thereof Analyzable: 14.6%                      Thereof Colexifying: 27.4%

Thereof by Contiguity: 10.5%                Thereof by Similarity: 26.6%

Recurrent associated meanings: fog/mist, sky, smoke, day, steam, rain, black/dark, cataract of eye, skin/leather, white, wind, faeces

33 sampled languages colexify ‘cloud’ (generally ignoring in the ensuing discussion if sources indicate that the relevant terms denote specific types of cloud unless this seems relevant) and ‘fog/mist’ (see § 6.3. for discussion of possible influence of environmental conditions, section 25 for analyzable terms for ‘fog/mist’ on the basis of ‘cloud,’ and Buck 1949: 65 for this pattern, which is pervasive by semantic shift in Indo-European). These are Bakueri, Ngambay, Dongolese Nubian, Yoruba, Anggor, Baruya, Buin, Kwoma (colexifying also ‘white,’ for which compare Bororo *baru kigadu* ‘sky clear/white’), Meyah, Ngaanyatjarra, Rotokas, Waris, Kosarek Yale, Basque, Bezhta, Sora, Haida, Lesser Antillean

Creole French, Nez Perce (by a lexical affix), Xicotepec de Juárez Totonac, Arabela, Aymara, Maxakalí, Piro, Ancash and Imbabura Quechua, Tehuelche, Yanomámi, Hani, Lenakel, Manange, Sedang, and Takia (note also the similarity between Embera *hīrará* ‘cloud’ and *hīrār’ā* ‘fog, mist’). Colexification with ‘smoke’ is found in nine languages, Nez Perce (again by the lexical affix also colexifying ‘fog’), Arabela, Bora, Cavineña (also colexifying ‘smell’), Cayapa (by the term *ñivijcha*, perhaps containing *ñi* ‘fire, flame, seed’ and *vijcha* ‘difference in height’), Maxakalí, and Tsafiki, and in Sedang by the analyzable term *kia hia* ‘ghost light.weight,’ colexifying also ‘air;’ moreover, Buin colexifies ‘white cloud’ and ‘white smoke’ (as well as ‘to be smoking tobacco, be affected by smoke’) specifically. Colexification with ‘steam, vapor’ is less frequent (attested in five languages, Anggor, Arabela, Bora, Cayapa (again by the term *ñivijcha*), and Tsafiki; note also the similarity between Koyraboro Senni *duule ~ duula* ‘cloud’ and *dullu* ‘smoke, steam’).

As for complex terms, Cashinahua has *nai kuin* ‘sky smoke’ for ‘cloud’ and Hupda has *j’ik kudu* ‘smoke group’ with the meaning “cloud, cloud of smoke/fog.” A further common cross-linguistic association is that with ‘sky’ (see also Buck 1949: 65 for evidence from Latvian, Welsh, and Breton). It occurs by colexification in Efik (by the analyzable term *ikpa’enyōñ*, with *ik’pa* meaning ‘animal skin’ inter alia and *en’yōñ* ‘sky, heavens;’ the association with ‘skin’ is mirrored by colexification, also with other related meanings, in Toba, and compare also Kolyma Yukaghir *ninqar* and *qar* ‘skin’), Lavukaleve, possibly in Nunggubuyu, Bezhta, dialectally in Sora, Upper Chehalis, Cheyenne (the meaning ‘sky’ is considered insecure by lexicographers in this language), Comanche, Kiowa, Lakota (here also colexifying ‘heaven’), Nez Perce, Miskito, and Rama (by the analyzable term *núnik kás* ‘sun/day meat;’ note in this context that Rao *grača* ‘cloud’ appears to contain *gra* ‘sun’ and that Yei *miramakər* seems to contain *makər* ‘flesh’). As for complex terms, alongside the one in Cashinahua already mentioned above, Efik has *idiök’ enyōñ /idiök’ en’yōñ/* ‘bad/unpleasant-sky,’ Yoruba *àwọ̀sánmà /awọ̀-sánmà/* ‘color/image-sky’ and *ojú-sanmà* ‘eye-sky’ (alongside *ojú òrun* ‘eye heaven’), Dadibi *bulu gi dagadu* ‘taro.type garden sky’ (denoting an “alto-cumulus cloud formation” specifically), Ket *espul /ēs-hul/* ‘sky-mound,’ Central Yup’ik *qilaggluk /qilak-rrluk/* ‘sky-one.that.has.departed.from.its.natural.state,’ Bororo *baru kigadu* ‘sky clear/white’ (meaning ‘white cloud’), Guaraní *arai /ára-i/* ‘sky/day-DIM’ (*ara* also has further meanings, and the term is said to be a contraction of *ára* in the sense of ‘weather’ and *vai* ‘bad’), and Jarawara *neme sabi* ‘sky/up be.foamy.’ Furthermore, the association is realized by noun class alternation in Swahili, and there is a semianalyzable term in Wayampi, which also denotes a magical plant that makes clouds come. Alongside Guaraní and Rama, which have an association between ‘cloud’ and ‘day’ by virtue of colexifying ‘sky’ and ‘day,’ there are also other cases with this pattern. Blackfoot *soksistsikó* is analyzable as */sok-iksistsikó/* ‘above-be.day,’ and in Hawaiian and Samoan, *ao* may refer to both ‘cloud’ and ‘day’ (in Hawaiian, also to the ‘dawn’ and other concepts and in Samoan also to the ‘head’), but notably not to the ‘sky.’ Furthermore, Badaga *karē* is glossed as “overcast sky, clouds hovering overhead.”

Occasionally, as already seen in the example from Kyaka and Bororo, terms make reference to either the white or black/dark color of clouds. This distinction is not exclu-



sively found in motivated terms; for instance, Berik has *anis* 'white cloud' and *umiyan* 'black cloud,' whereas Chayahuita *huiriro'tě* and *yararo'tě*, 'white cloud' and 'black cloud' respectively, contain the respective color terms. Further complex terms involving either constituents meaning 'black' or 'dark' include Mali *achēluinggi*, derived from *chluing* 'black,' Sahu *lobi-lobi'i*, reduplicated from *lobi'i* 'dark,' and Lesser Antillean Creole French *an mas nwe* 'in bulk dark,' while Piro colexifies 'cloud' and 'darkness' directly.

A linguistic pattern peculiar to New Guinea are terms literally translatable as 'wind faeces': Toaripi has *mea-e* 'wind/weather faeces' and Takia *tim-tae-n* 'wind-faeces-3SG.' Four sampled languages, Mbun, Nunggubuyu, Yir Yoront, and Cahuilla, colexify '(rain)cloud' and 'rain' (Yir Yoront also 'rainbow serpent,' note that this is a typical instance of actual/potential-polysemy common in Australia). Furthermore, in Khoekhoe, a term for 'cloud' in the Nama dialect, *Inanu-s*, shares the root *Inanu-* with *Inanu-b* 'rain, thunderstorm,' and the meaning is differentiated by the alternation of the nominal suffix. Kwoma *wayi keyi* is analyzable as 'rain make.mark.' Two languages, Hausa and Khalkha, extend a term for 'cloud' metaphorically to 'cataract of the eye' (the Hausa term with this pattern denotes 'light fleecy clouds' specifically and also bears the meaning 'cuttle-fish,' while the Khalkha term may also refer to a "cloud-like pattern or design").

Further patterns include: a further Hausa term denotes 'small clouds' and "a perfectly circular calabash;" this term, *san doki*, goes back etymologically to *sawun doki*, literally 'horse's footprint.' *Girgije*, yet another Hausa term, denotes "[l]argish clouds in wet season" as well as 'to shake off,' said of garment, plumage, and other things. In Khoekhoe, *lâub ~ lâus* is derived from *lâu-* 'expect, await' (perhaps because the presence of clouds arouse the expectation of rain? In fact, *lâus* also means 'expectation, hope'). Another register-specific Khoekhoe term for 'cloud' and 'rain,' *!hūduru-el*, is literally 'ground-mouse,' since the first drops of rain "make marks like footprints of mice." The Noni term *kembew* contains the noun class marker *ke-* and is otherwise identical with *mbew* 'beside, near,' while Dongolese Nubian *nīčč(i)* 'cloud, mist' is historically derived from *nīg* 'to sew' and hence also denotes the 'action of sewing;' the extension to 'cloud, mist' is explained in the source by them apparently veiling the environment like cloth. The Kwoma term *apoduwan keyi* (*apoduwan* 'crested or crowned pigeon;' *keyi* 'make mark') denotes 'small black clouds,' Rotokas *rukuta* may contain *ruku* meaning 'hinder, hold back, block' among other things, and Sko *a* also denotes the 'blackpalm' inter alia. Sahu *samasamama* 'fleecy clouds' is apparently formed by partial reduplication from *samama* 'crocodile.' Basque *laino* figuratively may also assume the meaning 'confusion,' and Japanese *kumo* also means 'spider.' The Pawnee term *cka'u* also means 'face,' while Tuscarora *urá?θeh* also means 'spleen.' Another term conceptualizing clouds using an internal organ as the vehicle is found in South America, namely Cubeo, which has *oco-penibo* 'water-liver.' San Lucas Quiaviní Zapotec *x:ca'āi* seems derived from *ca'āi* 'very early in the morning' by means of the nominal possession prefix *x:-*. Aguaruna *yujagkim* is also the name of a valley, and the Bororo term *tugu* can refer to 'cloud' as well as to a 'shadow' and 'suspicion, decision' alongside other things. Huambisa *yuragkim* contains *yurag* 'fruit,' but is not further analyzable, while Maxakalí *gōy* denotes 'cloud,' 'smoke,' as well as 'flood.' The Piro term *mko* is not only used for clouds, but also for 'tufts' or 'clusters' more generally, Ancash Quechua *puyu* also means 'moth,'

while Fijian *ō* is also used figuratively with the meaning ‘sign, trace’ (alongside being an interjection in ceremonial settings). Hawaiian *ōpua* ‘puffy, cumulus or billowy cloud, cloud bank’ is derived from *pua* which inter alia can mean ‘flower, blossom,’ but also ‘to issue, come forth, emerge,’ said of smoke, wind, speech and colors. Rotuman *aoga* is also the gerundive form of *ao* ‘to seek,’ and Sedang *xok* is also used with the meaning ‘garbage.’

### 13. *The Coal*

Representation: 93

Motivated: 29.3%

Thereof Analyzable: 12.8%

Thereof Colexifying: 16.4%

Thereof by Contiguity: 6%

Thereof by Similarity: 3.6%

Recurrent associated meanings: embers, fire, black, ashes, burn/burnt, stone, soot

Frequently, languages colexify ‘coal’ and ‘embers,’ as do Koyraboro Senni, Noni, Swahili, Yoruba, Gurindji, Kyaka, Ngaanyatjarra, Nunggubuyu, Upper Chehalis, Cheyenne, Chickasaw, Wappo, San Lucas Quiaviní Zapotec, Copainalá Zoque, Arabela, Carib, Guaraní, Toba, Wichí, Yanomámi, Kapingamarangi (by the analyzable term *ma-lala* ‘RESULTATIVE-heat.up’ which also denotes “wood chips made when using adze”), Manange, and White Hmong. In addition, Wayampi has the complex term *apĩ-nge* ‘embers-PAST’ and Kiowa has *’eip-k’q̣uɣɨ* ‘live.coal-black.’ In fact, it is frequently the case that terms for ‘coal’ make reference to its black color, as in Blackfoot *sik-óóhkotok* ‘black-stone’ or Biloxi *pě’xěno”n’ sũpi’*, containing *pe’ti ~ pětí’ ~ pět* ‘fire’ and *sũpi’ ~ sũpi ~ sũp* ‘be black.’ Further languages in which one of the constituents is the color term for ‘black’ or means ‘dark’ are found in Kwoma (*keyihapa now* ‘black color/earth.pigment,’ which in fact also means “black paint, black pigment used to make paint” as well as “object burnt by fire” generally), Nuuchahnulth (*tumiš /tum-řš/* ‘dark-consume’), Cavineña (*etiru-sehueda* ‘burnt-black,’ which is also capable of referring to a type of tree and the black fruit of a tree), Cubeo (*ñemichichi /ñemico-chichi/* ‘black-CLASS.SCALE-LIKE.OBJECT’), and Yanomámi (*ishiishi*, reduplicated from *ishi* ‘black, burnt’). For Tasmanian, Plomley (1976: 180) mentions the term *ly.hooner veene*, consisting of *ly.hooner* ‘black’ and *veene* ‘wood;’ note also the apparent lexical relationship between Cahuilla *tũl* ‘coal’ and *tũl-nek ~ tũl-ek* ‘black’ and between Bwe Karen *phe θàrò* ‘coal’ and *khi θàrú* ‘be very dark.’ In addition, Badaga and Tuscarora colexify the meanings (Badaga also with ‘vegetables, curry’ and “envy, jealousy, grudge”), and a semianalyzable term for ‘coal’ involving the word for ‘black’ might be present in Abzakh Adyghe.

Alongside Biloxi, terms for coals in which ‘fire’ acts as a contiguity anchor are found also in a number of other languages, and the respective terms are frequently metaphor-driven. Efik *u-kañ makara* is analyzable as ‘NMLZ-fire european’ and denotes the “charcoal used by smiths,” Yoruba *eyin iná* is literally ‘egg fire’ (*eyin* alone can also be used to refer to ‘coal’), Meyah *mah ofóm* is analyzable as ‘fire ripe,’ Waris *suwesumbul ~ suwumbul* as */suwe-sumbul/* ‘fire-earthworm,’ and Abipón *nkaate l-ahak* as ‘be.burning 3sg-wood’ (Macaguán colexifies ‘coal’ with ‘firewood,’ and semianalyzable terms with constituents meaning ‘tree’ or ‘wood’ are also found in Mbum and Embera). Carib *wa’to-po*, which colexifies ‘embers,’ is derived from *wa’to* ‘fire,’ Jarawara *yifo witi* is analyzable as

‘fire/firewood edge/nose/button’ (this term also colexifies the meanings ‘brand,’ ‘torch,’ and ‘fire’ itself), Rama *abung kúng* as ‘fire louse,’ Tsafiki *nin calá* as ‘fire silver/money,’ and Tetun *ahi-knaar* as ‘fire-work.’ There are semianalyzable terms featuring a constituent meaning ‘fire’ in Noni, Sko, Upper Chehalis, Bora, Guaraní, Huambisa, and Rama, and in San Lucas Quiaviní Zapotec, *gyih* is glossed as “fire: esp., coals, embers.”

Alongside Blackfoot, in which the term for ‘coal’ contains that for ‘stone,’ Carrier *tšě-tizken* is analyzable as ‘stones-burn,’ and Great Andamanese *taili châpa* consists of *taili* ‘stone’ and *châpa* ‘firewood.’ Khoekhoe *lui-nom-n* is analyzable as ‘stone-produce.charcoal-NMLZ,’ and a semianalyzable term where one of the constituents is ‘stone’ is found in Huambisa. Furthermore, six sampled languages, Burarra, Kwoma, Yir Yoront, Khalkha, Jarawara, and Sedang colexify ‘coal’ and ‘(hot) ashes,’ and in Yaqui *maatum* ‘coal’ consists of *maatu* ‘ashes’ and the plural marker *-(i)m* (a semianalyzable term for ‘coal’ containing the word for ‘ashes’ is also found in Bwe Karen). Two languages, Rendille and Yanomámi, colexify ‘coal’ and ‘soot.’

Further associations include: Buli *kaali* also means ‘to rub with hand,’ Ngambay *kər* also means ‘forest, bus,’ inter alia, and *kúl* also ‘to cook, prepare food.’ Gurindji *kunyini* also means ‘hearth,’ while Muna *gheo ~ ghio* also means ‘to cover with charcoal’ as well as “dry and cold (of corn, coconuts etc.).” Kyaka *langa* also means ‘careless, lax, impetuous,’ and Toaripi *aro* is also the name of the frigate bird (which is mostly of a deep black color) as well as a kind of squid. Nez Perce *símux* is also a man’s name, Wintu *kal* also means ‘to stare, gaze, intense visual contact, glow,’ ‘receive intense visual impression/sensation’ and ‘feather,’ and Central Yup’ik *qetek* is the name of the “underground tuber of the *mare’s-tail plant* (*Hipperus vulgaris*) or *horsetail plant* (*Equisetum arvense*)” and also means “piece of coal, from its similar, black appearance,” while Lesser Antillean Creole French *chebon ~ chabon* also means ‘carbuncle.’ Aguaruna *kayúshik* might contain *kayút* ‘to smoke,’ and the Yanomámi term *ōshōhōre* also denotes the ‘burnt bone of a dead person,’ Bislama *kol* also means ‘cold, cool,’ (due to phonological collapse of Eng. *cold* and *coal*), ‘sticky, adherent,’ and “greeting, call (on radio),” and Hani *meiq* also ‘to teach.’ Hawaiian colexifies ‘coal’ with ‘bituminous,’ Lenakel *nouanamkīlu* contains *noua* ‘fruit,’ and Rotuman *mahala* may also refer to ‘cinders,’ ‘coke,’ and ‘clack lead.’

#### 14. The Coast

Representation: 59%

Motivated: 53%

Thereof Analyzable: 41.2%

Thereof Colexifying: 12.4%

Thereof by Contiguity: 17.2%

Thereof by Similarity: 22.7%

Recurrent associated meanings: edge/end/border/point/limit, sea, water, side, river/stream, land, hem, proximity

Clearly the most frequent lexico-semantic association for the ‘coast’ (or ‘shore,’ ‘beach’) is that with meanings such as ‘edge,’ ‘end,’ ‘border,’ ‘point,’ or ‘limit’ (see Buck 1949: 31-32 for Indo-European evidence, where this association is well-attested). This may be by colexification, as in Noni, Kyaka, Bezhta (where the relevant term also means ‘land’ in

general, as is the case in Dongolese Nubian; there is a semianalyzable term where this is the meaning of the identifiable constituent in Rotuman), Khalkha, Sora (colexifying also 'line'), Santiago Mexquititlan Otomí, Wintu, Aguaruna, Bororo, Chayahuita, Sáliba, Toba, Tehuelche (where the relevant term *-ork'en* is derived from *-or* 'around' and also means 'body'), and Yanomámi, but occurs more frequently by morphologically complex terms, with the second constituent being 'river,' 'sea,' or more generally 'water,' as in Buin *atigaminno* 'sea-edge' or Dadibi *ái bogabadu* /*ái bogabadu*/ 'water border.' Further languages with terms of this type are Yoruba (where 'edge' is colexified with 'ear'), Sahu, Toaripi (where 'point' is colexified with 'corner,' 'peak, summit, topmost point'), Basque, Khalkha (by another term than the colexifying one listed above), Laz, Chickasaw, Lesser Antillean Creole French, Pawnee, Yaqui, Bora (where 'edge' is colexified with 'side,' note in this context that Gurindji *pirntiwirti* is reduplicated from the base *pirnti* 'side,' and that Aguaruna, Bororo, and Toba colexify 'coast' and 'side'), Guaraní, Hupda, Hawaiian, Malagasy, White Hmong, and Yay. All in all, the association is found in 31 sampled languages. Variants of the association are found in Efik, where *adaña*, which also denotes a 'morning ebb' among fishermen and 'mark, average' generally is derived from *daña* 'to fix boundaries, fix extremes,' Ket, where 'land' rather than 'edge' or 'border' is the meaning of the relevant constituent (*sest ba'ŋ*, analyzable as /*sēs-d-ba'ŋ*/ 'river-poss-land' and *ul-baŋ-t*, possibly 'water-land-NMLZ') and Wichí, where *tewuk lhip* is literally 'river part' (Blackfoot has *iisóítahtaa* /*iso-niitahtaa*/ 'on.a.horizontal.surface-river'). There are also structurally different complex terms in which one of the constituents is 'sea.' One such term is Khoekhoe *huri-ammi* 'sea-deep.furrow' (*ammi* also means 'beak'), another one is Huambisa *nayantsa pakari*, containing *nayants* 'sea, ocean' and *paka* 'plain.' Yir Yoront, where there is also a complex term of the redundant type *yuwl-ther* 'sea/seashore-edge,' Abzakh Adyghe, Badaga, and Yuki colexify 'coast' and 'sea,' the latter language by the complex term *ʔuk hoʔ* 'water big,' while Mali colexifies 'coast' and 'salt water.' Semianalyzable terms one of the constituents of which is 'sea' are found in Huambisa, Rama, and Piro. Furthermore, Miskito and Rama, spoken in close proximity and in contact, share terms for the 'coast' that are strikingly similar: Miskito has *kabo lama* 'sea proximity' (though *lama* also means 'breast' inter alia) and Rama *táuli shá suka* and *táuli ska* ~ *táuli ská*, containing *táuli* 'sea' and *súka* 'next.' Other structures are also found for complex terms on the basis of 'water': here, the cross-linguistic variation includes terms such as Kanuri *cí nǝ̀-bè* 'mouth water-of,' Carrier *yǝthúpa*, containing *yǝ* 'sky' and *thú* 'water' and the Kashaya terms *qʰatow*, analyzable as /*ʔahqʰa=tow*/ 'water=at' and *qʰami lahwál*, containing *ahqʰa* 'water' and *lahwál* 'across.' There is a semianalyzable term where the identifiable constituent is 'water' in Manange. Khalkha and Yanomámi colexify 'coast' with 'hem,' the relevant Khalkha term *købege(n)* ~ *købyge* appears to be derived from *købe-* 'tack, baste in sewing,' suggesting that 'hem' is the original meaning.

Other associations include: Efik *usuk* and *ikpekhe* are derived by a nominalizing prefix from the verbs *suk* 'bring down' (this term may also refer to "[c]ountries beyond, out of the river, e.g. England" and occasionally to the 'south') and *pekhe* 'divide' (this term also may refer to a 'division,' 'curtain,' and 'apartment' inter alia). The variant *gǝ̃r* of Dongolese Nubian *gǝ̃r* ~ *nǝ̃r* ~ *ǝ̃r* also means 'to crush (grain)' and 'cave' (the latter reading is

due to borrowing from Arabic). Swahili has *pwa-ni* 'ebb-LOC,' Yoruba *àgbègbè* also means 'neighborhood, vicinity' and 'region,' and Anggor colexifies 'coast' with 'sand.' Berik has *fi eyep* 'salt around' (note also Noni *ngem* (yoo), derived from a verb meaning 'to surround, encircle' by a noun class prefix; yoo is 'water'). Burarra *gochilawa* "coastal area, on land or out on sea" is derived from *gochila* 'abdomen,' while Nunggubuyu *madhalag* has a narrow reading 'beach, coastline' but may more generally refer to the 'coastal region' (the same is true of Basque *kostalde*). Basque *costa* also means 'to dock, berth.' Lesser Antillean Creole French *lakòt* can also refer to "places away from the city" generally, Nez Perce *ʔalláy* denotes "a low place, bottom of valley" generally and also "the downriver region (toward the Pacific Coast); downward; toward the Pacific Coast." Santiago Mexquititlan Otomí *ñāni* also means 'to defend.' Pawnee has the term *huukaahaaru*?, containing *huuka*(wi)- 'along a stream course' and *haar* 'place;' this term also means 'dry river or lake bed.' Somewhat opaque semantically is Tuscarora *kəʔ tiwakəhyaʔnáʔnyeʔ*, which is based on the verbal root *-akəhyaT-* 'to extend from.' Wintu *ʔeʔ* also denotes the 'edge of fence or earthlodge,' 'at the edge, marginally' and 'all over, toward,' *ʔoq* also means 'bar, rocky bar' and 'along,' while, similarly, Central Yup'ik has *canineq*, which contains *cani* 'area beside' and the postbase (see § 4.4.2.) *-neq* 'area of.' This term also has the meaning 'area.' Arabela has a semianalyzable term containing an element meaning 'lagoon, pool,' Bororo *iba* is also the name of the place outside the village where fishermen gather before hunting or fishing to evoke certain spirits and to distribute tasks. Carib colexifies 'coast' with 'wharf,' Imbabura Quechua with 'beside,' Wayampi inter alia with 'lip,' and Yanomámi with 'periphery' and 'contour.' Great Andamanese *tôtgôra* and *îgora* appear to be derived from *gôra* 'be durable,' and Hawaiian *makālae* 'beach, shore, coast near a point' contains *lae* 'point.'

#### 15. *The Dew*

Representation: 77%

Motivated: 30%

Thereof Analyzable: 16.3%

Thereof Colexifying: 14.1%

Thereof by Contiguity: 6.0%

Thereof by Similarity: 8.5%

Recurrent associated meanings: water, fog, cold/frost, hoarfrost, urine, snow, wet/moist, star, ice, drizzle, night, wind

In eight sampled languages, Burarra, Rotokas (by the term *vusiva*, for which compare *vusi* 'burst forth, erupt, break out?'), San Mateo del Mar Huave, Quileute, Arabela, Rama, Hawaiian, and Bwe Karen, 'dew' is colexified with 'fog, mist' (in Burarra, the colexified meaning is 'heavy fog' more specifically, and Rama also colexifies 'smoke'). Similarly, Wintu colexifies 'dewy' with 'misty.' In San Mateo del Mar Huave, the relevant term colexifying 'dew' and 'fog' is *qjmiüc naquind* 'fall cold.' This is, however, not the only language in which 'dew' is lexically associated with meanings like 'cold,' 'cool' or 'frost.' Badaga, Greek, Comanche, Itzaj, Hawaiian, and Tetun colexify one of these meanings with 'dew;' in Rotuman, 'dew' is *matit toka* /matiti koka/ 'cold be.calm/be.still' (this term also means "cool night breeze or air") and Ngaanyatjarra *nyinnga kumpu* is analyzable as 'frost urine' (note also that Carrier *terhzö*, containing *terh* 'above,' also means "hoar-frost on trees," a pattern

of colexification also encountered in Buli, Cavineña, and Tehuelche). Furthermore, Tehuelche and Hawaiian colexify ‘dew’ with ‘ice,’ and Buli, Badaga, Cavineña, and again Hawaiian colexify ‘dew’ and ‘snow.’ These associations are of little surprise when one bears in mind that dew is caused by cooling down of the air which causes atmospheric water to condense. Another subsidiary explanation may be recent semantic extension of ‘dew’-words to ‘ice’ or ‘snow’ in regions where these states of water do not naturally, or only seldomly occur, as is suggested in the consulted source for the colexification of ‘dew’ and ‘snow’ in Buli.

There are many sampled language where ‘dew’ is expressed by a complex term with one of the constituents meaning ‘water.’ However, there is, with the exception of Nez Perce, no language in the sample that colexifies these two meanings (the closest to this is Lenakel, which uses the same term for ‘dew’ and “water on the ground after rain”). Terms with such a structure include Khoekhoe *lgâ-lgam-mi* ~ *lgâ-lgam-mi* ‘grass-water-3SG.MASC,’ Abzakh Adyghe *weseps* /we-se-psə/ ‘sky-??-water/liquid,’ Nivkh *fi-sax* ‘bore-water,’ Itzaj *p’uj-ja* ‘swarm-water,’ Central Yup’ik *merr’aq* /meq-rraq/ ‘fresh.water-a.little.bit.of’ (this term also means ‘holy water’), Miskito *diwas laya* ‘wind.from.land.or.east liquid’ (note also the colexification of ‘wind, storm’ in Kyaka and “gentle land breeze, as of West Hawai’i” in Hawaiian), Imbabura Quechua *yaku wiki* ‘water drop,’ Wayampi *iapi* /ii-api/ ‘water/river-source’ (colexifying ‘dew’ with ‘valley’ and ‘spring’), Yanomámi *heuheu*, presumably containing *he* ‘head, upper or lower extremity’ and *u* ‘liquid’ (an alternative term is *hemaahu*, presumably containing *maa* ‘rain’ alongside *he*), Bislama *wata blong naet* ‘water of night’ (in Muna, ‘dew, expose to dew, put something out overnight to cool’ is *aloma* ~ *alo*, with *alo* also meaning ‘night, evening;’ this association is also not surprising given that the air cools down in the night and thus dew is typically found in the evening, night or morning), and Tetun *mahon-been* ‘shadow-liquid.’ Kiliwa has a derived term, and semianalyzable terms where the identifiable constituent is ‘water’ are found in Kosarek Yale, Ket, and Guaraní. Moreover, there is the idiolectal term *ok-yobi-* ‘water-be.spring-NMLZ’ in Chickasaw.

In five languages, namely Angkor (perhaps), Upper Chehalis, Ineseño Chumash, Kiliwa, and Hani, the term for ‘dew’ is lexically related to a word for ‘wet’ and ‘moist.’ In Upper Chehalis, the word for ‘dew’ is *sáx<sup>w</sup>sáx<sup>w</sup>*, reduplicated from *sáx<sup>w</sup>* ‘wet, damp,’ and in Ineseño Chumash, *spuyan* is derived from *puy* ‘being moist.’ In Kiliwa, the association is by colexification, and in the rest of the mentioned languages the relevant terms are semianalyzable only. There is a notable area in the Northwest Amazon region where metaphorical terms for ‘dew’ are found, involving either ‘star’ (note again that the cause of dew is condensation of atmospheric water at night), ‘urine,’ or both as a source concepts. In Bora, the word for ‘dew’ is *mñícúru ní-jpa*, probably analyzable as ‘star urine-CL.liquid,’ in Arabela, it is *riya shaaca* ‘star:PL urine,’ in Hupda, it is called *wædhəm’æh nəcáw* ‘star saliva.’ While in these particular configurations, the association is clearly an areal phenomenon on a small scale (see Epps 2007: 285 and Zamponi 2009: 590 for data from Maipure specifically, but note also that Farr 2001: 126, citing T. and C. Weber in personal communication, reports *fial ok* ‘star water’ for Bimin in New Guinea), the association with ‘urine’ is also found elsewhere: Kaluli *eleya: ba:* contains *elé* ‘moon’ and *ba:* ‘urine’ and Ngaanyatjarra *nyinnga*

*kumpu*, as mentioned above, is analyzable as 'frost urine.' Finally, Wintu *p<sup>h</sup>On* and Arabela *shushiquia* also may refer to a 'drizzle.'

Further associations include: Efik *mbara* also denotes "drops of rain hanging on the bushes after a shower" and Hausa *ra'ba* is also the name of "[a] disease of goats and sheep in which 'bags' of fluid are found in the abdomen (like highly magnified dewdrops)," inter alia. Khoekhoe *laub* may be related to *lau*, a verb meaning 'to trickle, purl' inter alia, which also yields the formally identical *laub* 'spring, fountain.' Ngambay *tàal* is also the name of the 'jackal,' the Rotokas terms *uriteva* and *visiva* also mean 'manna' (presumably used in Bible translations), and Sko *kú* also means 'special armband.' Yir Yoront *kal* is also a kinship term while *kalvmr* also denotes the 'galah,' also known as rose-breasted cockatoo. Badaga colexifies 'dew' also with 'hail' and 'cold season,' and Bezhta *xida* also means 'snot.' Kiowa *ʔ'da* contains *ʔ* 'smoke, misty rain' and *da* 'to be' (and is incidentally identical segmentally with the term for 'tree, wood' and 'plant'), and Santiago Mexquititlan Otomí *xa* also means 'to mow.' San Lucas Quiaviní Zapotec *go'ohp* also means "dumb, stupid, slow-moving." Aguaruna *jegashík* appears to contain *jéga* 'house,' while Chayahuita *poroncayo* possibly contains *cayo* 'egg.' In Yanomámi *hemaahu*, the first constituent might be *he* 'head,' but all these associations in languages of South America seem quite dubious semantically. Hupda *paç nán* is analyzable as 'sky grease,' Fijian *yau* inter alia also means 'to carry, to bring,' Hawaiian colexifies 'dew' with 'to blow' (as of a breeze), 'to hit,' and other things, Bwe Karen colexifies 'gallbladder,' and Yay 'stripe,' 'design,' and 'to write' inter alia.

#### 16. *The Dust*

Representation: 92%

Motivated: 54.2%

Thereof Analyzable: 12.4%      Thereof Colexifying: 41.8%

Thereof by Contiguity: 5.6%      Thereof by Similarity: 19.3%

Recurrent associated meanings: powder/grit, land/soil/earth, dirt/rubbish/garbage, ashes, smoke, sand, flour, steam, fog/mist, spray of sea, pollen, gray, sawdust, insect droppings

Notable are the lexical associations found between 'dust' and some aerosols. Frequently, languages colexify 'dust' and 'smoke' (found in Bezhta, Wintu, Carib, Hupda, Ancash Quechua, Bislama, where also 'cigarette' and other meanings are colexified, Fijian, Hawaiian, and Sedang, see also Buck 1949: 18 for evidence from Indo-European) or have complex terms for 'dust' on the basis of 'smoke,' a pattern found in Guaraní, Tsafiki and Tetun, where, for instance, 'dust' is *rai-bolon* 'earth-smoke.' In a subset of these languages, Carib, Guaraní, Hupda, Ancash Quechua, Tsafiki, and Fijian, an association with 'steam' is also found, due to colexification of 'smoke' and 'steam;' in addition, in Hupda, the redundant complex term *məy j'k* 'house smoke/steam/fog/dust' is available to single out the 'dust'-reading of *j'k*. Tsafiki also betrays an association with 'cloud' and furthermore, four sampled languages, Abzakh Adyghe, Wintu, Hupda, and Hawaiian, also have an association between 'dust' and 'fog.'

However, these are neither the only nor the most frequent cross-linguistic associations. In 26 sampled languages, associations with ‘dirt,’ ‘rubbish,’ or ‘garbage’ are found, by colexification in Buli, Efik, Burarra, Kwoma, Kyaka, Ngaanyatjarra, Southeastern Tasmanian, Badaga, Chickasaw, Highland Chontal, Ineseño Chumash, Kiowa, Nuuchahnulth, Oneida, Pawnee, Wintu, Central Yup’ik, Aguaruna, Arabela (colexifying ‘street dirt’ specifically), Aymara, Bororo, Cayapa (by a semianalyzable term containing *pe* ‘faeces’), Huambisa, Hani, and Sedang. Nuuchahnulth, in addition, has the derived term *mačq-mis* ‘dirty-NMLZ,’ and Khalkha has the analyzable term *boy say* ‘filth/rubbish/sweepings egg.of.insect’ with the meaning “dirt, dust, esp. in a house, refuse, rubbish,” which is mirrored by a Rotokas term colexifying ‘dust’ and ‘insect droppings.’

22 sampled languages, without the emergence of any clear areal pattern in their distribution, colexify ‘ashes’ and ‘dust’ (again, see Buck 1949: 18 for parallel Indo-European data). These are Efik, Berik, Burarra (where the relevant term is glossed as “little particles, as ashes and sand mixed where a fire has been burning” more specifically), Lavukaleve, Mali, Nunggubuyu, Rotokas, Basque, Chukchi, Biloxi, Carrier, Upper Chehalis (colexifying ‘cold ash’ specifically), Highland Chontal, Lesser Antillean Creole French (by the analyzable term *sann cho* ‘sand fire), Lake Miwok, Oneida, Tuscarora, Wintu, Yuki, Cavineña, Hani, and Bwe Karen. Furthermore, one of the Buli terms for ‘dust’ is *tanbuulum* /taungbuulum/ ‘sand/soil-powdered’ (which also colexifies “a special kind of fine white clay (chalk?) that is eaten (sucked) by many Ghanaian people”), in Kaluli, ‘dust’ is *sa-mu* ‘sand-ash,’ in Bislama, it is *asis blong graon* ‘ashes of ground/soil,’ and in Mandarin there is the redundant compound *hui1-chen2* ‘ashes-dust.’ Equally frequent are languages in which there is a term for ‘dust’ which also has the more general meaning of ‘powder’ or ‘grit,’ as in Kwoma, Kyaka, Muna, Basque, Chukchi, Greek, Kildin Saami, Welsh, Biloxi, Highland Chontal, Lesser Antillean Creole French, San Lucas Quiaviní Zapotec, Arabela, Cavineña, Cubeo, Embera, Lengua, Piro, Ancash Quechua, Toba, Malagasy, Manange, Mandarin, Rotuman, and Samoan, (and several Indo-European languages, Buck 1949: 18). Furthermore, in Maxakalí, ‘dust’ is *hãpkônôn*, putatively analyzable as /hap-kōnnôn/ ‘roast-powder,’ Bwe Karen has *(de)gəmu* ‘(thing-)be.powdered’ and *ha phí mu* ‘earth husk be.pulverized,’ and White Hmong *hmoov av* ‘powder soil.’ The association with ‘sand’ is not only found in the languages mentioned so far. By colexification, it also occurs in Efik (also with ‘clay’), Burarra (here the relevant term *baluk* is said to mean “little particles, as ashes and sand mixed where a fire has been burning”), Ngaanyatjarra, Toaripi (by a semianalyzable term containing a constituent meaning ‘earth’), Khalkha, Chickasaw, and Arabela (where the relevant term also means ‘beach’ and ‘sandpit’), and Tehuelche *te:mxen* ‘dust’ contains *te:m* ‘soil, sand, wandering dune.’ Seven sampled languages, Kildin Saami, Biloxi, Tuscarora, Sáliba, Wayampi, Manange, and Mandarin, colexify ‘dust’ and ‘flour’ (Wayampi also ‘spot of reddish paint on face’), while in Mbum, ‘dust’ is *sômn-jâl* ‘flour-soil’ and in Imbabura Quechua *allpa jaku* ‘land flour.’ As has become clear from several of the examples cited above, ‘land,’ ‘soil,’ or ‘earth’ are frequently involved as a contiguity anchor in complex terms for ‘dust.’ Further terms of this kind that do not fall in any of the categories of semantic associations discussed so far include Yir Yoront *larrpown*, which is analyzable into the words for ‘ground, earth, soil dirt,’ *larr*, and ‘strike,’ *pow*, (both con-



stituents also colexify other meanings) followed by a noun theme formative, Kiliwa *?-mat-sa?* ‘DN-earth-grease/fat,’ Santiago Mexquititlan Otomí *fonthai*, analyzable as /font'i-hai/ ‘spread.out-earth,’ and Miskito *tasba yuyu* ‘earth/soil/place grain/little.particle,’ while Ineseño Chumash *šupšup* ‘dust’ appears to be reduplicated from *šup* ‘earth.’ Furthermore, Bakuri, Efik, Dongolese Nubian, Yoruba, Dadibi, Ngaanyatjarra, Southeastern Tasmanian, Nuuchahnulth, Oneida, Wintu, Arabela, Aymara, Ancash Quechua, Toba, and Hawaiian feature terms that colexify ‘dust’ with ‘earth,’ ‘soil,’ or specific types thereof (Yoruba also colexifies ‘mud’). Arabela *joojojua* colexifies ‘starch’ and ‘milk powder.’

Two sampled languages, Tuscarora and Yuki, associate a color term for ‘gray’ with ‘dust,’ Tuscarora by colexification and Yuki by the analyzable term *poʔot̚el*, containing *poʔ-* ‘burn’ and *poʔ-* ‘gray’ (these terms both also colexify ‘ashes’); a semianalyzable term containing an element meaning ‘gray’ is also found in Upper Chehalis. Languages which employ the same term for ‘dust’ and ‘spray of sea’ are Khoekhoe, Fijian, and Hawaiian (Hawaiian also colexifies ‘foam’ inter alia, and Fijian also “the foam at the front of a swiftly moving canoe” specifically). Three languages of Oceania, Kyaka, Kosarek Yale, and Hawaiian have terms which colexify ‘dust’ and ‘pollen’ (Kyaka *kuku* in general may refer to a “bit, flake, tiny fragment, crumb, dust, talc, powder” and also denotes a kind of tree). Finally, Kildin Saami *puenn* also denotes ‘sawdust’ – a parallel is found in Chayahuita- and ‘metal filings’ specifically.

Other associations include the following: Efik *obu* is derived from *bu* ‘to rot, corrupt,’ *ntõñ* also means “dimness of vision as if a mist were before the eyes” and denotes a type of spearmint, *m'bio*, derived from *bi'o* ‘to cut,’ also means ‘chaff, trash, sweepings’ and ‘mote,’ and *n'tan* also denotes a plant, “the flowers of which eject a fine dust when touched.” Hausa *k'ura* also dialectally means “[a]ny drinking water in which nothing has been admixed” and ‘urine,’ Khoekhoe *tsarab* also means “dirt dust of cloud” and ‘smog,’ while Rendille *malálwa* can mean, alongside “fine dust found in certain places,” also ‘sandstorm’ and ‘duststorm,’ ‘dust cloud’ is also colexified with ‘dust’ itself in Gurindji, and in Upper Chehalis *słasóq* contains a root meaning ‘dusty’ and ‘dust flying.’ Dadibi *pusugu bage* may contain *bage* ‘mixed group,’ Kwoma *hapasen* might contain *hapa* ‘bone,’ while Meyah *ofóu* is peculiar in that it is glossed meaning ‘egg,’ ‘meaning,’ and ‘dust,’ which is likely due to homonymy. Ngaanyatjarra *kuwiri* is also used with reference to “a place where a kangaroo has lain.” Rotokas has terms for ‘dust’ which are based on verbal roots for ‘leave’ and ‘arrive’ respectively: *kavurao*, containing *kavu* ‘leave,’ and *poupou* which apparently is reduplicated from *pou* ‘arrive.’ Kosarek Yale *lolonga* also means ‘chips, chip-pings’ and “remains of a fire or an earth-oven” and *ubdoba* also “powder on butterfly wings.” Ket colexifies ‘dust’ with ‘mould.’ Welsh *lluwch* also means ‘spray’ and ‘snowdrift.’ A semianalyzable Upper Chehalis term for ‘dust’ colexifying ‘cold ashes’ specifically appears to contain a constituent meaning ‘inside.’ Comanche *huhkup̣̥ ~ huukup̣̥* appears to contain *huh ~ huup̣̥ ~ huuhp̣̥ ~ huh-* ‘tree, wood,’ while Tuscarora *awéheʔ*, alongside ‘flour,’ has also been extended semantically to mean ‘gunpowder,’ and San Lucas Quiaviní Zapotec *deheh* to ‘washing detergent.’ Wintu *sud*, as a verb, also means ‘to cure with smoke’ and ‘disinfect,’ and *bukul* ‘ashes, dust, dirt, soil’ contains *buk-* ‘dark.’ Central Yup'ik *caarrluk* contains *ca-* ‘something’ and the postbase (see § 4.4.2) *-rrluk* ‘one that has departed from

its natural state.’ Cashinahua *kudu* can also be used to refer to the color ‘cream,’ Jarawara *hobokori* appears to be derived from *hoboko*, the name of a tree species, and Rama *úng ulúng* contains *úng* ‘pot’ (and colexifies ‘pot dirt’). Hani *gaoqbee* also means ‘weeds.’ Great Andamanese *êrl’ôtbûbut* is analyzable as /êrl-l’-ôtbûbut/ ‘place-??-3SG.POSS.BODY.PART-SOOT,’ and there is another not otherwise analyzable term containing *êr* ‘place.’ Hawaiian *lele-huna* ~ *lele-hune* “[f]ine windblown rain spray, dust, mist; to fall as fine rain” is analyzable ‘fly-particle,’ and *lepo*, another term for ‘dirt’ in the language, can also be used to refer to ‘excrements’ and ‘silt’ inter alia. Kapingamarangi *bopobo*, colexifying ‘decayed’ and ‘mildew,’ is derived from *bobo*, meaning ‘rotten, decayed, old.’ Tetun colexifies ‘dust’ with “to crumble, fall apart, smash” among other meanings, and, finally, Yay *pun*<sup>5</sup> can also mean ‘manure, fertilizer.’

### 17. The Eclipse

Representation: 41%

Motivated: 72.2%

Thereof Analyzable: 63.1%

Thereof Colexifying: 9.2%

Thereof by Contiguity: 12.8%

Thereof by Similarity: 56.9%

Recurrent associated meanings: sun, moon, darkness/darken, die/kill, eat, catch, be ill, disappear, numb, swallow, go into

Terms for this concept are, where data are available, in the overwhelming majority of languages expressed by morphologically complex metaphor-driven terms, with the meanings ‘sun’ and/or ‘moon’ acting most often as contiguity anchors. The most commonly recurring pattern features lexical associations with either ‘to die’ or ‘to kill.’ This is found in Chickasaw, Central Yup’ik, Bororo, Chayahuita, Sáliba, Kapingamarangi, Lenakel (where ‘to die’ is colexified with ‘be ill,’ San Lucas Quiaviní Zapotec colexifies ‘be in eclipse’ with ‘be ill, have menstrual period,’ and Ancash Quechua has *rupay qishyan* /*rupay qishyay-n*/ ‘sun become.ill-3SG’), and Tetun by morphologically complex terms (for instance, Tetun *loro-mate* ‘sun-die’), and by colexification in Central Yup’ik which has the complex term *iraluq nata-luq* ‘moon/month die-??,’ but also colexifies the relevant meanings in the verbal root *nala-*, which can either mean ‘to die,’ ‘to become numb,’ and ‘to be eclipsed’ as well as convey other meanings. Furthermore, Kiliwa has *t+hiy=h+nyuu* ‘OBJ+spirit=3+kill’ and Samoan *gase-toto* ‘be.numb/be.dead-blood’ (note the parallelism in the association between numbness and death in Central Yup’ik and Samoan).

Another pattern that is common in particular in languages of Southeast Asia is complex terms for ‘eclipse’ that are based on the respective words for ‘eat’ plus the name of an animal. Thus, in Yay, an ‘eclipse’ is called *baaŋ<sup>2</sup> kun<sup>1</sup> duan<sup>1</sup>* ‘flying.squirrel eat/drink month’ or *baaŋ<sup>2</sup> kun<sup>1</sup> taan<sup>1</sup>van<sup>4</sup>* ‘flying.squirrel eat/drink sun,’ in Sedang, a lunar eclipse is called *kau ka khê* ‘kau.fish eat moon’ and a solar eclipse *koxê ka hâi* ‘centipede eat sun,’ while in Hani, the ‘eclipse’ is called *ba’la neivq-kee qaq* ‘moon spirit-dog eat’ or *naolma neivq-kee qaq* ‘sun spirit-dog eat,’ the “spirit-dog” is said in the consulted source to figure in a traditional story. Furthermore, Sora and Upper Chehalis have complex terms for the eclipse involving verbs meaning ‘to swallow’ and ‘to eat’ respectively (the association with

'swallow' is diachronically detectable in Haida as well), and the Xicotepec de Juárez Totonac term *hua'can*, which is combined with the respective terms for 'sun' and 'moon' to refer to a solar or lunar eclipse respectively, contains *hua* 'eat, drink.' Some degree of areal convergence may be assumed to be responsible for terms for the 'eclipse' found in three African languages which involve a verb meaning 'to catch' (Hausa *rana tā kama wata* 'sun 3SG do catch moon,' Mbum *séù/sèsèi ṅgbánà hánà séù/sèsèi* 'sun/moon catch other sun/moon' and Yoruba *ímúṣòkùkùn*, analyzable as /ì-mú-ṣe-òkùkùn/ 'NMLZ-catch-do-darkness,' similarly Khalkha has *nara(n) barixu* and *sara barixu*, containing the respective term for 'sun' and 'moon' and *bari-* meaning 'to hold, grasp, seize, catch' inter alia). In fact, terms based on the meanings 'darkness' or 'darken,' as in Yoruba, are relatively frequent cross-linguistically, occurring either by terms of the lexical or derived type also in Khoekhoe, Dadibi, Basque, Arabela, and Fijian (for instance, Basque has *ilun-aldi* 'darkness-time' and *itzal-aldi* 'shade-time,' the first term may also refer to 'darkening' or, figuratively, 'sadness' or 'confusion,' and the second also to 'darkness' or 'a period of darkness'). Particularly noteworthy is also the Dadibi term *gilga ge begelama hulia saidao*, as it seems to be based on the idea that the sun turns over to the other side and thus does not shine anymore, the literal translation offered by lexicographers is "sun having turned over, (it) had gotten darkness." The association with 'darkness' or 'darken' is also found by colexification, namely in Cheyenne, Piro, and Hawaiian. In three sampled languages, Oneida, Cavineña and Guaraní, terms for 'eclipse' involve verbs meaning 'to disappear' and the secondary meanings 'to get lost' in Cavineña and 'flee' in Guaraní. In two languages of the Americas, Carrier and Tuscarora, the 'eclipse'-terms are based on a verb meaning 'to go into' (*sa yā-te-aiḥ* 'sun sky-recess-pocket-get.in' and *θkà'yēh* respectively which is analyzable as /čì-ka-yē-h/ 'REPETITIVE-3SG.INDEF.AGENT-go.into-HAB,' note in this context also Nez Perce *?ipné'culeylekse* which revolves around the verb *leylé'k* 'to move into a hole').

Of course, there is also a number of other metaphor-driven analyzable terms for either lunar or solar eclipse in which either 'sun' or 'moon' act as contiguity anchor. These include Hausa *wata yā yi zazza'bi* 'moon 3SG.MASC do feverishness' and *wata yā yi mashashara* 'moon 3SG.MASC do small.pox/feverishness,' Noni *diuu bo kpwee dwee lo* 'sun ?? moon bridge take.by.force' and *diuu ε baṅ kpwee* 'sun to cover moon' for 'lunar eclipse' and *diuu bo kpwee ε tasen e mvunsheej*, which involves the words for 'sun' and 'moon' as well as *tasen* 'to meet' for the 'solar eclipse,' Nivkh *k'jeṇ mu yr* 'sun boat time,' and Kaingang *mīg tȳ kysā/rā mǎn* 'panther with sun/moon carry.' Quileute *xitlītssil pitītscho?* is not amenable to a precise morphological analysis on the basis of the source, but the literal meaning is given as "a monster bit away chunks of the moon," and Toba *qayāpo'oguet na l'edaxa aṇi nala* contains 'edaxa' 'light, flame' and *nala* 'sun.' In addition, Great Andamanese *lajabaginga*, to which either *bôdo* 'sun' or *ôgar* 'moon' is added to refer to 'solar eclipse' and 'lunar eclipse' respectively, might contain *jabagi* 'to damage.' Semianalyzable terms on the basis of 'sun' or 'moon' are found in Ineseño Chumash and Cavineña.

Further isolated associations include: Buli *yesinta chaab joka* contains *yesinta* 'shadows' (the literal translation offered by the lexicographer is "shadows entering each other"), while Efik *erinime* is derived from *nime* meaning 'to extinguish, eclipse' inter alia, and indeed can also refer to 'extinguishing,' 'extinction,' or 'vanishing.' Katcha *thigirono*

*ma there* and *thigirono ma ndhinaia* are presumably related to *thigidono* ‘silence’ (*there* is ‘moon’ and *ndhinaia* appears to be a variant of *ndhanaya* ‘sun’). Welsh *diffyg* means ‘lack’ and *rhagori ar* is literally ‘excel on.’ Wintu *čil* may be related to a word for ‘bird’ and ‘bear’ of the same shape. San Lucas Quiaviní Zapotec *rda’au lohoh* contains *rda’au* ‘get shut, be stuck’ and *lohoh* ‘mouth, in front of, on top of.’ Bora *íjyúnubááve* contains *íjyúnu* ‘become night,’ and there is also a semianalyzable term containing an element meaning ‘be night’ in Blackfoot. Lesser Antillean Creole French *ékrips* is also used to refer to a ‘loss of brilliance.’

#### 18. *The Egg*

Representation: 97%

Motivated: 26.2%

Thereof Analyzable: 5.2%

Thereof Colexifying: 21.0%

Thereof by Contiguity: 1.4%

Thereof by Similarity: 1.3%

Recurrent associated meanings: testicles, seed, fruit, child, hen/chicken/bird, nut, spawn, ovum/zygote, male genitals

Terms for the ‘egg’ of a bird or a chicken specifically are cross-linguistically frequently associated with other smallish round objects, such as ‘fruit,’ ‘nut,’ ‘seed,’ or ‘stone’ in a mixture of analyzable and colexifying terms. In seven sampled languages, Buin, Toaripi, Kosarek Yale, Comanche, Hawaiian, Samoan, and Takia, ‘egg’ is lexically associated with ‘fruit.’ As becomes clear from the above list, this pattern is particularly frequent in languages of the broader New Guinea area (see McElhanon and Voorhoeve 1970: 29, Laycock 1970: 1141-1142). The association is mostly realized by colexification, except for Toaripi (*ori fare* ‘bird fruit’) and Comanche (*kokorá?a pokopi* ‘chicken fruit’). Also in New Guinea, colexification with ‘nut’ is found, specifically in Dadibi (indeed, the relevant term is also glossed as ‘small object’), Kaluli, Kyaka, and Takia (note that Takia belongs to both groups, and indeed, it is the same lexeme that can refer to either ‘egg,’ ‘fruit,’ or ‘nut’). The same areal hotspot in New Guinea is discernible for an association with ‘seed,’ and also here, Takia, as well as Kyaka, participate in the pattern by virtue of the same term having a wide semantic range. Other languages in which ‘seed’ and ‘egg’ are colexified are Buin, Rotokas, Kosarek Yale, Abzakh Adyghe (inter alia), Wintu (where *lu* also means to ‘stab, pierce, poke’), and Hawaiian, while Baruya colexifies ‘sprout,’ ‘shoots,’ and ‘seedling.’ In addition, in Kiliwa, ‘egg’ is *xma?=yit-y* ‘hen=seed-ATT.’

A pattern common in North America is to have terms for ‘egg’ based on ‘child’ and/or ‘daughter/son’ more specifically. This is found in Japanese (*tama-go* ‘ball-child’), Chickasaw (*akankoshi*’, analyzable as /akanka’-oshi/ ‘chicken-son;’ this term in fact also denotes a ‘chick’), Comanche (*kokorá?a atüap̣* ‘chicken child’), and by colexification in Kiowa, San Lucas Quiaviní Zapotec, and, outside of the Americas, in Kosarek Yale and Hawaiian. A somewhat similar term is found in Abipón, where *tetarik l-kaoe-te* is analyzable as ‘hen POSS.INDEF/3SG-work/creature-PL.’ In general, the cross-linguistic evidence reveals that terms for ‘egg’ are sometimes morphologically complex with ‘bird’ acting as a contiguity anchor (a semianalyzable term is in addition found in Kemtuik). In contrast, there is

no language in the sample in which the reverse situation holds, i.e. in which the word for 'bird' is secondary to that for 'egg' (compare section 6; interestingly, unlike the situation observable for the meanings 'bee' and 'honey,' there is also no language which colexifies 'bird' and 'egg'), and this provides cross-linguistic support for the assumption that in Proto-Indo-European, the word for 'egg,' *\*h<sub>2</sub>ō(w)i-om* in Mallory and Adams's (2006: 143) reconstruction, is derived from that of 'bird,' *\*h<sub>2</sub>ewei-*, and not the other way around, see Schindler (1969) for more detailed discussion.

Furthermore, Koyraboro Senni colexifies 'egg' and 'stone' (although the relevant term assumes the meaning 'egg' only in compounds), and in Santiago Mexquititlan Otomí *mādo* 'egg' contains *do* 'stone' (see Buck 1949: 256 on this association in Baltic languages). A very common and well-known metaphorical transfer pattern connects the meanings 'egg' and 'testicle' (also in evidence in Indo-European according to Buck 1949: 256). This is found in the sample in as many as 19 languages by colexification (and in nine by complex terms, see section 142), namely Efik, Ngaanyatjarra, Carrier, Highland Chontal, Itzaj, Santiago Mexquititlan Otomí, Pawnee (with slight deviations in form), Pipil (Cuisnahuat dialect), Tuscarora, Bororo, Miskito, Piro, Ancash Quechua, where the relevant term also colexifies 'scrotum' and 'rattle (snake),' Imbabura Quechua, Rama, Tsafiki, Yanomámi, Bwe Karen, and Lenakel. In One and Sedang, moreover, 'male genitals' rather than 'testicles' specifically are colexified with 'egg.'

Four languages, Efik, Kyaka, Sora, and Wintu, colexify 'egg' and 'spawn,' and two languages, Basque and Hawaiian, colexify 'egg' and 'ovum' or 'zygote.'

Other associations include: the Efik term *nse* also denotes "a granule of any thing" as well as the 'foetus in the womb,' while *ñqua* is also "the general name of a bead." Ngambay *kàbbè* is also the name of a large kind of tree, Kaluli colexifies 'egg' with 'center,' and Kyaka *kapa* can also refer, inter alia, to the 'core of something,' a 'larva,' a 'friend,' or 'iron.' There are two further languages in the sample where the term for egg has other meanings which do not seem to be in any substantial relation with 'egg': Meyah *ofou* also means 'meaning' and 'dust,' and Sentani *do* which also means 'man' and 'name.' The Kosarek Yale term *wana* also means 'heart' and "flower-stalk," Yir Yoront colexifies 'egg' with 'bud,' and Sora *adre:ŋ ~ arre:n* is derived from *arre:-* 'lo lay eggs' (this may also be the situation recoverable etymologically for Basque; compare also Hani *alwuv ~ hawuv* and *wuv* 'to hatch,' Yay *ʔan' cay²* 'CLASS.THING lay egg,' and for Jarawara *hife/hifene* 'egg' *hifa* 'brood, lay eggs'; Itzaj colexifies 'egg' and 'to lay egg.'). Abzakh Adyghe *-č'e* also means 'to grow' inter alia. Highland Chontal colexifies 'egg' with 'cocoon,' and Kiowa is unusual in apparently employing the same term for 'child,' 'egg,' and 'semen,' *iH* (there is also the term *iH-ṭṭe* 'child/egg/semen-white' to single out the 'egg'-reading), and the general meaning of the Wintu *pe:λ* in *pe:λabuhabe* is stated to be "rounded, smooth and shiny," and the term denoting 'eggs' containing it is also used to refer to 'marbles,' 'apples,' 'watermelons' and moreover means 'baldheaded.' Cashinahua *bachi* can also refer to a 'piece of clothing' as well as to a 'wasp nest' and a 'blackhead.' Cayapa *napipu* may contain *pu* 'thorn, thin bone.' Cubeo *jīdū* (analyzable as *ji*, which bears the lexical semantic meaning 'egg' and the classifier *-dū* for roundish three-dimensional objects) can also be used to refer to any kind of protuberance in general. Maxakalí *xu'uk* has a short form *xuk*; there is a verb of that form

meaning “to carry in a bag or sack, to be pregnant.” Ancash Quechua *lluntu* can also be used with reference to ‘hail’ (although this term is said to be specific to child language), and Toba *lco’oue* is derived from a verb meaning ‘to give birth.’ The correct analysis (if any) of Tsafiki *pi’poca* is not straightforward; it might consist of *pi* ‘water, liquid’ and *poca* ‘cane of the Guadua bamboo.’ Hawaiian *hua* has a wide range of potential referents, including alongside those already discussed also ‘tuber,’ ‘produce, yield’ alongside ‘round objects’ in general. Kapingamarangi *ngogo* can also mean ‘brain’ and ‘zero,’ probably both because of the similarity concerning the roundish appearance, while Rotuman is unique in colexifying the meanings ‘egg’ and ‘chrysalis.’ Samoan *fua* also means ‘flower, bloom’ as well as ‘products,’ similar to the cognate Hawaiian term, and Takia *patu-* may also mean ‘shell’ and ‘back.’ Yay *cay*<sup>2</sup> is also a verb meaning ‘to lay eggs.’

#### 19. *The Embers*

Representation: 64%

Motivated: 55.9%

Thereof Analyzable: 26.2%

Thereof Colexifying: 30.2%

Thereof by Contiguity: 16.5%

Thereof by Similarity: 6.9%

Recurrent associated meanings: coal, fire, flame, ashes, burn, hot/heat/warm, spark, cinders, wood/firewood, small, red, grain

Rather than being based on similarity, terms for the ‘embers’ (or ‘live coal’) are most frequently associated lexically with meanings from the same domain by contiguity, that is, meanings being related in some way to ‘fire.’ Six sampled languages, Kosarek Yale, Khalkha, Carrier, Kiliwa, Cubeo, and Ancash Quechua colexify ‘embers’ and ‘(hot) ashes’ (Carrier also ‘dust’), and in two languages, the term for ‘embers’ is related by word-formation to that for ‘ashes’ (Mandarin *hui1-jin4* ‘ashes-what.is.left.from.fire’ and Yoruba *eéru gbígbóná* ‘ashes hot’). Still more frequent, however, is an association with ‘coal,’ and this pattern also occurs formally by both colexification and morphological analyzable terms. Among the colexifying languages are Koyraboro Senni, Noni, Swahili, Yoruba, Gurindji, Kyaka, Ngaanyatjarra, Nunggubuyu (the term is also glossed as “heart of fire”), Upper Chehalis, Cheyenne, Chickasaw, Wappo, San Lucas Quiaviní Zapotec, Copainalá Zoque, Arabela, Carib (by the term *wa’topo*, derived from *wa’to* ‘fire’), Guaraní, Toba, Wichí, Yanomámi, Kapingamarangi (by the term *malala*, derived from the verb *lala* ‘to heat up’ and colexifying “wood chips made when using adze”), Manange, and White Hmong, and it may be present diachronically in Ket. Where the association is by analyzable terms, the additional constituent is often also from the same semantic field. For instance, Buli has *kaala bolim* ‘coal:PL fire/light’ (‘fire’ is also present in Lesser Antillean Creole French *ti mòso chabon difé* ‘small piece coal fire,’ which also has *sann cho* ‘sand hot/burning’), Basque *ikatz bizi* and *ikatz gorri* ‘coal live’ and ‘coal red,’ Bororo *eradu uru ~ joradu uru* ‘coal heat,’ Miskito *kwasku lakni* ‘coal flame,’ Kaingang *pránh gru* ‘coal flame/burning,’ and Fijian *qilaiso sa waqa tū* ‘charcoal PART be.burning stand’ (‘burn’ is one of the additional constituents also in Sko, compare also Lesser Antillean Creole French *sann cho* ‘sand hot/burning’ and Abipón *leer-gRaie*, which consists of grammatical material that is built around the root *eerg-* ‘burn,

sparkle;’ this term colexifies ‘nettle’). Similar to the associations with ‘red,’ Vietnamese has *than hồng* ‘coal pink,’ and in addition, there are semianalyzable terms one of the constituents of which is ‘coal’ in Yuki, Miskito, and Sáliba. All of the above associations also recur in different configurations. Meanings like ‘heat’ and ‘warmth’ are also relatively frequently associated in the languages of the sample, occurring alongside the cases of Yoruba, Lesser Antillean Creole French, Bororo, and Kapingamarangi that were already mentioned also in Kiliwa (*kw+pal*, which also colexifies ‘ashes,’ is analyzable as ‘PERF+hot;’ note the parallel to the Carib term where the morpheme acting to derive the complex term makes reference to the past), and Great Andamanese (*arpjil’igûya*, presumable containing *pj* ‘hair, feather’ alongside *ûya* ‘warm’). A term for ‘red’ is alongside Basque also one of the constituent of the relevant Biloxi term, where *pě’xěno”ni’ tcti’* contains *pe’ti* ~ *pěti* ~ *pět* ‘fire’ and *tcti* ‘red.’ Alongside Sko, Lesser Antillean Creole French, Kaingang, and Fijian, ‘burn’ is also the associated meaning as revealed by analyzable terms in Japanese (by a derived term) and Bwe Karen (*mi-ú* ‘fire-burn/catch.fire’); *tequeyashi* ‘burning’ is furthermore one of the constituents of Arabela *tequeyashi nootunenu* ‘embers’ and Wayampi *api* contains *api-* ‘burn down.’ Furthermore Hupda *teg hō* ‘wood burn’ colexifies ‘fire,’ ‘flame’ and ‘embers,’ as does Badaga *kiccu* (the term also means ‘conflagration,’ ‘bonfire’ and, figuratively, ‘jealousy’ inter alia), while Buli colexifies ‘embers’ and ‘flame’ (but not ‘fire’), Cashinahua and Piro colexify ‘embers’ and ‘flame,’ and Aymara also colexifies all three meanings, but by two different terms. ‘Wood’ or ‘firewood’ are also sometimes encountered as constituents in complex terms, occurring, alongside the special case of Hupda, in Efik (*ñkpri ifia ikañ* ‘small:PL firewood fire’) and Sora (*əra:de’todən* /ə’ra:-de-tod-ən/ ‘wood-??-fire-N.SFX’). Four languages, Basque, Khalkha, Haida, and Tuscarora colexify the meanings ‘embers’ and ‘spark’ (in Basque, the relevant term can also mean ‘spunk, pep’ colloquially and in Khalkha it also means “slendour, grandeur, glory, energy” and ‘spirit’). In two sampled languages of the Americas, Santiago Mexquititlan Otomí and Imbabura Quechua, there are parallel terms that involve a metaphorical transfer from ‘grain’ to ‘embers’ by analyzable terms with ‘fire’ acting as a contiguity anchor. The Otomí term is *děspi* /*dě-tsibi*/ ‘fire-grain’ and the Imbabura Quechua one, *nina muru*, is also analyzable as ‘fire grain.’ As already mentioned above, ‘fire’ in general is unsurprisingly the most “popular” contiguity anchor cross-linguistically. Buli *bolim ngiak* might be analyzable as ‘fire origin’ or as ‘fire liquid’ (analysis is considered insecure by the lexicographer), Kosarek Yale *auk wana* is analyzable as ‘fire fruit/seed/egg,’ and semianalyzable terms where ‘fire’ is one of the constituents are found in Ngambay, Noni, Kaluli, Toaripi, Kosarek Yale, Cheyenne, Nez Perce, Pipil, Chayahuita, Guaraní, and Yay. Alongside colexification in Hupda, Badaga, and Aymara, San Lucas Quiaviní Zapotec *gyih* is glossed as “fire: esp., coals, embers.” Finally, Cubeo *ūacovū* may contain *ūaco* ‘cinder field,’ and Lengua and Tetun in fact colexify ‘embers’ with ‘cinders.’

Other associations include: Rendille is unique in using the same word for ‘embers’ and ‘testicle,’ *jiláh*. Yoruba *oguna şuşu* contains *şuşu* ‘greatly,’ and the relevant Angkor term may also refer to glowing metal. Kyaka *langa* also means ‘careless, lax, impetuous.’ The Mali term *sachongini* is derived from *sachon* ‘eye’ by means of the diminutive singular suffix *-ini*. Abzakh Adyghe *meś’ez’ek* contains *ž’ek*, ‘ember, surrounding of fire,’ which in

turn contains *k(e)* ‘surrounding, depression.’ Badaga *ganda* also means “sandalwood paste.” Welsh *marwor* appears to be derived from *marw* ‘to die,’ and *marwydos*, another Welsh term, seems to be related to the verb as well. Aymara *k’qja* also denotes a ‘loud color’ as well as ‘high temperature.’ The meaning potential of Tuscarora *uč’l’reh* includes “candle, ember, flash of light, lamp, light, spark, taper.” Central Yup’ik has *cupun* ~ *cup’un* /*cupe-n*/ ‘blow-device.for’ (this term also means ‘straw’ and ‘rifle’).

#### 20. *The Estuary*

Representation: 32%

Motivated: 49.6%

Thereof Analyzable: 36.5%

Thereof Colexifying: 15.2%

Thereof by Contiguity: 9.9%

Thereof by Similarity: 27%

Recurrent associated meanings: mouth, river/stream, opening, water, foot/leg, branch, last, sea, flow out/go out

The ‘estuary’ is a meaning more often expressed by terms with similarity rather than contiguity as the underlying semantic relation; in either case, the relevant terms are mostly analyzable. The most frequent metaphor-driven pattern is the association with ‘mouth,’ occurring in eleven of the sampled languages, namely Efik, Khoekhoe, Muna, Toaripi, Abzakh Adyghe, Welsh, Upper Chehalis, Cheyenne, Nez Perce, Fijian and Mandarin; moreover, there is a semianalyzable term in Lesser Antillean Creole French. In Efik, Khoekhoe, Toaripi, Abzakh Adyghe, Upper Chehalis, Nez Perce, and Mandarin, the association is by colexification (as well as in Latin according to Buck 1949: 228), in the remaining language it is realized by morphologically complex terms. In Welsh, Muna, and Fijian, the second meaning involved is ‘river,’ as in Muna *wobha-no laa* ‘mouth-POSS river,’ and in Cheyenne, the verb *ho’ome* meaning ‘be the confluence, be the mouth of a river’ perhaps contains *ho* ‘arrive’ and *óm* ‘mouth’ (in many of these languages there are also additionally colexified meanings peculiar to ‘mouth,’ such as ‘opening’ in general, ‘beak,’ etc., see § 6.2.3.2. for discussion). In addition, the Chayahuita term also appears to contain the word for ‘mouth’ and that for ‘to berth,’ and the Kolyma Yukaghir term *ajil* ‘opening, river-mouth’ derives diachronically from the same root as ‘mouth’ (note that Arabella *jiyacuaji* ‘estuary, entry to path’ contains *jiya* meaning ‘hole, earth’ inter alia). Somewhat parallel to the pattern of colexification in Kolyma Yukaghir, ‘opening’ is also colexified with ‘estuary’ and ‘mouth’ in Efik, Khoekhoe, Toaripi, Abzakh Adyghe, and Mandarin.

Another body-part metaphor that is however rarer is that based on ‘foot’ or ‘leg,’ which occurs in Muna and Tetun, in the former language by colexification, in the latter by the analyzable term *mota-ain* ‘river-leg/foot’ (note also that Haida *t’aa* also means ‘foot of trail’ and ‘foot of bed,’ compare the noun *st’a* ‘foot’). Chukchi colexifies ‘estuary’ with ‘throat,’ and Wayampi with ‘head’ inter alia. An association not based on a body-part that is still clearly metaphorical in nature is that with ‘branch’ occurring in Burarra (also with ‘arm, wing’) and Basque. In the latter language, the term for estuary is *itsas-adar* ‘sea-branch/horn,’ the association with ‘sea’ is also present by a derived term in Welsh.



Contiguity-based analyzable terms include ones based on a verb meaning ‘to flow out’ or ‘to go out’ in Chukchi and Nuuchahnulth, and the Kashaya and Piro terms for ‘estuary,’ which include a constituent part with the meaning ‘to fork off’ or ‘fork, intersection’ (compare Piro *skitha* ‘estuary’ and *ski* ‘base intersection, fork;’ this term also denotes a ‘confluence of rivers’). Obviously, analyzable terms frequently include constituents with the meaning ‘water’ or ‘river’ as contiguity-anchoring devices. Among these, ‘river’ is the most commonly used contiguous concept. Further terms of this kind include Kwoma *pa bogo* ‘river junction/bend,’ Sko *pá-long* ‘river-hole,’ Yanomámi *u paa thapi pa*, where *u* is ‘liquid, river,’ *thapi* is ‘further’ and *pa* ‘away’ (this term is said to mean ‘above the estuary’), and Bwe Karen *lâ-ká* ‘river-bottom.’ Moreover, Welsh colexifies ‘estuary, confluence’ and ‘stream’ directly. Terms with ‘water’ include Carrier *thû-ķě-tcě* ‘water-on-tail,’ Kashaya *šohq<sup>h</sup>awi*, analyzable as /ʔašo-ʔahq<sup>h</sup>a-wi/ ‘south-water-at’ (note for this term that *šohq<sup>h</sup>a* ‘south water’ is a site at the mouth of the Russian River in Northern California, so this term is more of a toponym rather than a general term for ‘estuary’), as well as the apparently cognate terms of Hawaiian *muli-wai* ‘after/last-water’ and Samoan *muli-vai* ‘be.last-water’ (*muli* has other meanings in both languages, in Samoan, inter alia ‘buttocks,’ so here a body-part metaphor might be underlying the conceptualization).

Further associations include the following; Buin *tope* also means ‘to dry out, wither,’ Burarra colexifies ‘mouth of river’ with ‘lower back, tail bone.’ This term also has the more general reading of “base or rear end of anything,” while the Yir Yoront term *ngopngrr* denotes the property ‘wide’ in a general sense, and, more particularly, a wide place in a river and an estuary. Another colexifying term is found in Khalkha, where *aday* also denotes any kind of ‘end’ inter alia. Lesser Antillean Creole French *labouchwi* also means ‘harbor,’ and Tuscarora *nyawé'kə?* is analyzable as /t-ya-ek-ə-ʔ/ ‘CISLOC-3SG.INDEF.PATIENT-liquid-fall-PUNCTUAL’ (presence of roots is sure, precise analysis otherwise inferred and perhaps erroneous).

## 21. The Feather

Representation: 95%

Motivated: 56.0%

Thereof Analyzable: 11.2%      Thereof Colexifying: 45.0%

Thereof by Contiguity: 10.4%      Thereof by Similarity: 40.6%

Recurrent associated meanings: hair, fur/wool, wing, leaf, bird, bristle, quill, pen, beard, needle of tree, wool, scale

Most commonly, terms for ‘feather’ are associated lexically with the meanings ‘hair’ and/or ‘fur’ and ‘wool,’ both by colexification as well as by morphologically complex terms, in which case most often ‘bird’ acts as a contiguity anchor. Colexification with ‘hair’ (without a difference made between ‘body hair’ and ‘head hair’ in the ensuing discussion) is found in as many as 41 sampled languages, namely Buli, Efik, Hausa, Dongolese Nubian, Anggor, Berik, Burarra, Kaluli, Kwoma, Kyaka, Meyah, Muna, Sahu, Sentani, Middle-Eastern and Southeastern Tasmanian, Toaripi, Kosarek Yale, Abzakh Adyghe (where the relevant term also means ‘tooth’ and ‘seed’), Kolyma Yukaghir, Biloxi, Chickasaw, San

Mateo del Mar Huave, Central Yup'ik, Copainalá Zoque, Abipón (colexifying also 'leather'), Arabela, Chayahuita, Guaraní, Lengua, Rama, Sáliba, Toba, Fijian, Great Andamanese, Hawaiian, Kapingamarangi, Bwe Karen (which also colexifies 'between' and "to be strong, forceful"), Lenakel, Malagasy (colexifying also 'moss'), Tetun, and Yay. Colexification with 'fur' or 'wool' is found in 39 languages, namely Dongolese Nubian, Swahili, Baruya, Burarra, Kaluli, Kwoma, Kyaka, Meyah, Muna, Nunggubuyu (where the reading 'feather' is rare), Sahu, Middle-Eastern and Southeastern Tasmanian, Toaripi, Kosarek Yale, Kolyma Yukaghir, Abzakh Adyghe, Chickasaw, San Mateo del Mar Huave, Central Yup'ik, Copainalá Zoque, Aguaruna, Arabela, Bora, Chayahuita, Guaraní, Huambisa, Kaingang, Lengua, Miskito, Rama, Toba, Tsafiki, Hani, Hawaiian, Bwe Karen, Lenakel, Tetun, White Hmong, and Yay. The two categories are not mutually exclusive, since a large number of languages does not distinguish between 'hair' and 'fur/wool' lexically. There are also complex terms such as Hupda *hūtē'h pā't* 'bird hair,' which are also found in Yei, Laz, Wichí, and Vietnamese, and thus obviously without any particular areal hotspot in a certain region of the world. Other complex terms where one of the constituents is 'hair' include Mali *chēsengvēs*, derived from *kēseng* 'hair,' Santiago Mexquititlan Otomí *xi'ni/xi-oni* 'body.hair-chicken' and Kapingamarangi *ngaa-hulu* 'movement-hair.' Complex terms on the basis of 'fur' or 'wool' are less frequent, and are only encountered in Chukchi (*yalya-rəyrəy* 'bird-fur') and Manange (*1jima-2m'wi* 'bird-fur'). Due to very vague semantics of the Kosarek Yale term for 'hair'/'feather' which also includes 'beard' and 'fur' in its semantic extension, an association with 'beard' is also diagnosed in that language, as is the case in Bororo. An association with 'scale' is found in Middle-Eastern and Southeastern Tasmanian and Bororo.

Other terms with 'bird' as a contiguity anchor include Sko *táng-hó* 'bird-undress' (colexifying 'bird's tail') and Bakueri *yali yá wūwa* 'leaf of fowl,' as well as the similar Toaripi *ori tolo* 'bird leaf' for "long feather, from wing or tail," and there are semianalyzable terms in Mbum and Kyaka. In fact, Bakueri and Toaripi are not the only languages in which an association between 'feather' and 'leaf' is found, although they are the only ones in the sample where it is realized by an analyzable term. By colexification, the association is also found in Burarra and Ngaanyatjarra (colexifying "broad leaf" more specifically), and in seven languages of the Americas, namely Chickasaw, Ineseño Chumash, Abipón, Cashinahua (colexifying also 'silver, money'), Jarawara, Lengua, and Toba. In Kolyma Yukaghir and Chickasaw, in addition, the relevant terms also denote the needles of a coniferous tree.

In 18 sampled languages, Kanuri, Yoruba, Gurindji, Lavukaleve, Ngaanyatjarra, Badaga, Greek, Sora, Acoma, Upper Chehalis, Highland Chontal, Nez Perce, Cayapa, Jarawara, Kaingang, Macaguán, Piro, and Bwe Karen, terms which colexify 'feather' and 'wing' exist, a relation also very common in Indo-European (Buck 1949: 246). In addition, Cubeo *cave-do* 'feather' is analyzable as 'wing-CLASS.LARGE.CYLINDRICAL.SLIM.AND.ACUTE.OBJECT.' Buli, Kwoma, Bora, and Rama also colexify 'feather' with 'bristle,' and Basque, Cheyenne, Lesser Antillean Creole French, and Hawaiian also with 'quill.' Finally, three languages, Basque, Nivkh, and Lesser Antillean Creole French colexify by (at least initial) functional contiguity 'feather' and 'pen' (Nivkh by further extension also 'pointed drill'). In Buin, *paru* 'feather, plumage' is glossed also as "(used for) flower," and in Nunggubuyu *-dhabag*

'feather' can also refer to the 'flowers' of certain plants (see Laycock 1970: 1145; 1975: 228); the relevant Buin term can also refer to a "tangle-net trap for phalangiers, flying-foxes, rats, and other small game."

Other associations include: Buli *kok* can also refer to a 'ghost,' Hausa *gashi* also means 'character' *inter alia*, and *jawaska* is also used with the meaning 'the habits and customs of a country.' Baruya features unrelated terms for feathers removed from the bird and feathers when they are on the bird. Muna *wulu* also denotes a kind of small bamboo and means 'to have a dessert' as a verb, while Ngaanyatjarra *nyarlpi* can also refer to 'playing cards.' In Yir Yoront, *marr* means 'feather' and 'wing-feather' specifically, and, by a metonymic transfer typical for the Australian language area, also denotes the 'red-winged parrot' (*Aprosmictus erythropterus*) which has a bright red (hence, "salient") wing feather. Rotokas *orupa* contains *oru* 'trim down, shave away,' Abzakh Adyghe *c(e)* may also refer to a fine substance, Basque *luma* can also refer to a 'snowflake' and *hegats* also to a 'fin' and 'eaves' (this term may be related diachronically to *hatz* 'finger'), Greek *fteró* also to a 'feather-duster' and 'mudguard,' Sora *bə'le:dan* also to "the plume worn by the Sora-men on a turban," and Cheyenne *mée'e* also means 'collarbone.' The Kiowa term, according to the consulted source, may be related to a lexeme with the meaning 'tree, wood, stick,' and Lake Miwok *pútte* also denotes "that part of the dance headpiece which fits on top of the head." Nez Perce colexifies '(large) feather, wing' with "eagle with white-tipped wings; golden eagle, *Aquila chrysaetos*," Wintu *kal* also means 'to stare, gaze, intense visual contact, glow,' 'receive intense visual impression/sensation' and 'coal' and *kalaq* also 'feather headdress,' while Central Yup'ik *melquq* is analyzable as /meqe-quq/ 'shed.hair-one.that.is' (compare perhaps also *buuni* 'feather' in the Yahi and Northern Yana dialects of Yana with *buu* 'to shed hair,' the association is however considered unsure in the consulted source). Bororo colexifies 'feather' with 'shell' of animals and insects as well as 'casing.' Toba *laue* also means 'tower, observation point.' Yanomámi *hakoraki* 'wing feathers' appears to contain *hako* 'shoulder.' Finally, Bislama colexifies 'feather' with 'pubic hair,' 'grass,' 'fern,' and 'mould,' Hawaiian with 'brush' *inter alia*, Hani with 'to be stuck onto, to smear,' Bwe Karen with 'to spread,' and Rotuman with 'eaves.' Finally, White Hmong *plaub* also means 'four.'

## 22. The Flame

Representation: 74%

Motivated: 63.2%

Thereof Analyzable: 36.3%

Thereof Colexifying: 27.8%

Thereof by Contiguity: 27.1%

Thereof by Similarity: 26.2%

Recurrent associated meanings: fire, tongue, light, burn, embers, spark, ray/beam of light, blaze, wood, lobster, flare, candle, knife blade

Terms for the 'flame' are frequently formally complex, with 'fire' typically being the contiguity anchor. The dominant pattern cross-linguistically is metaphorical in nature and involves a transfer from 'tongue' to 'flame.' This is very frequent, and is found in all areas of the world in a total of 24 languages (Efik, Hausa, Kanuri, Katcha, Mbum, Noni, Swahili, Yoruba, Baruya, Kyaka, Toaripi, Abzakh Adyghe, Bezhta, Kildin Saami, Sora, Upper Che-

halis, Highland Chontal, Ineseño Chumash, San Mateo del Mar Huave, Fijian, Lenakel, Malagasy, White Hmong, and Takia). Canonically, the terms are of the lexical type, as in Katcha *ngado m-isi* ‘tongue ??-fire,’ with the variation that in Hausa the association is by colexification, in Lenakel, the association is (“possibly,” according to the consulted source) realized by reduplication (*nam* ‘tongue,’ *namnam* ‘flame’), and in Abzakh Adyghe, the relevant term *tx’ebz* ~ *tx’abze* is analyzable as /tx°(e)-bze/ ‘catch-tongue’ rather than ‘fire-tongue.’ Note also Chayahuita *pën nënërinso*, containing alongside *pën* ‘fire’ *nënërin* meaning ‘to flame’ as well as “to stick out the tongue,” as well as Lake Miwok *létaw-*, which also means “to dart the tongue (said of a snake).” Other terms with ‘fire’ acting as a contiguity anchor are Buli *bolim ngiak*, which might be either analyzable as ‘fire origin’ or ‘fire liquid,’ Kanuri *kánnú cámbî* /kannu cambi/ ‘fire it.has.given.birth.to,’ Yoruba *ṣwṣṣ-iná* ‘hand/branch-fire,’ Japanese *honoo* /hi-no-ho/ ‘fire-GEN-ear,’ Carrier *khwen-tlek* ‘fire-uses.to.dart.out,’ Bora *cúújuwa péétene* /cúújuwa peéte-ne/ ‘fire light-CL.thing.or.action,’ Bororo *eru-gu* ~ *loru-gu* ‘fire-liquid/blood,’ Jarawara *yama hirini* ‘thing catch.fire,’ Miskito *pauta klauanka ba* ‘fire/firewood burning DEM,’ Wichí *itoj lhalh* ‘fire brightness,’ Hawaiian *lapa ahi* ‘ridge/slope fire,’ and Vietnamese *ngọn lửa* ‘peak fire,’ while Malagasy *lèla* in *lèlào* ‘flame’ colexifies ‘tongue and ‘blade’ (*àfo* is ‘fire’). There are semianalyzable terms where the identifiable constituent is ‘fire’ in Kosarek Yale, Waris, and Santiago Mexquititlan Otomí. Moreover, Sedang has *pla on* for ‘flame,’ with *on* meaning ‘fire’ and *pla* inter alia “to mediate, to reconcile, to break up a fight, to come between and stop a quarrel.” However, similar to the association with ‘blade’ above, there is also *plá* ‘knife blade’ which may be the constituent rather than *pla*.

In addition, Chukchi *ḡəlenyəlet* /ḡəlet-ḡəlet/ ‘burn-warm.up,’ Hupda *teg hō* ‘wood burn,’ and Yanomámi *koā wake* /kōā wake/ ‘wood red’ colexify ‘flame’ and ‘fire,’ and this is also the case for simplex terms in Dongolese Nubian, Berik, Badaga, Ket, Khalkha, Comanche, Nez Perce, San Lucas Quiavini Zapotec, Arabela, Cayapa, Tehuelche and Mandarin, so there are in total fifteen languages with colexifying terms. Alongside Chukchi and Hupda, in eight further languages, Upper Chehalis, Cheyenne, Xicotepec de Juárez Totonac, Kaingang, Miskito, Kapingamarangi, Rotuman, and Samoan, associations with a verb meaning ‘to burn’ are encountered, in Kaingang, Kapingamarangi, and Rotuman by colexification, in the others by morphologically complex terms: Upper Chehalis *skáwmitn* ~ *skáwitn* contain *káw-* ‘burn.’ Cheyenne has *exo’áséotse* /éxo’asé-otse/ ‘burn-become,’ Xicotepec de Juárez Totonac *lamaná* is analyzable as /lama-na/ ‘burn-AGT,’ Miskito *pauta klauanka ba* as ‘fire burning DEM,’ and Samoan has a derived term (‘burn’ is also a common association in Indo-European, Buck 1949: 72). Similarly, Sahu colexifies ‘high flames’ specifically with ‘to burn fiercely.’ Thirteen sampled languages, Dongolese Nubian, Buin, Gurindji, Ngaanyatjarra, Nunggubuyu, Yir Yoront, Sora, Wintu, Abipón, Bora, Bororo, Lengua, and Miskito have terms for the meaning ‘flame’ which betray a lexico-semantic association with meanings such as ‘(fire) light’ or ‘to light’ occurring in a mixture of colexification and analyzable terms of the derived and lexical type. Colexification occurs in Dongolese Nubian, Gurindji, Ngaanyatjarra, Nunggubuyu, Rotokas, Yir Yoront, Wintu, Abipón, Bororo, Guaraní, Lengua, and Hani (Gurindji also colexifies ‘torch,’ Ngaanyatjarra also ‘electricity,’ and Yir Yoront also ‘light-source’ and ‘flame color’;

Bororo *eru-gu* ~ *loru-gu* is analyzable as 'fire-liquid/blood' as mentioned above, but colexifies 'light' in addition, and the same is true of Abipón *l-irie-Ra*, which is analyzable as 'POSS.INDEF/3SG-ignite-ABSTR'). Sora *gən'e:m-ən* ~ '*ge:m-ən* is derived from the verb *gən'e:m-* ~ '*ge:m-* 'to light,' and Bora has, as already mentioned, *cúújuwa péétene*, which is analyzable as /*cúújuwa* *peéte-ne*/ 'fire light-CL.thing.or.action.'

In Swahili, Tuscarora, and Bororo, 'flame' and 'ray,' 'beam of light' are colexified (and in Tuscarora in addition 'halo'). Further, Meyah *éisa* is glossed as 'flame, shine' (probably it is a verb). In four languages, colexification of 'flame' with 'spark' is found. In two of them, Abzakh Adyghe and Chayahuita, the respective terms are also formally analyzable and fall into the category of terms with constituent elements meaning 'fire' and 'tongue,' while in San Mateo del Mar Huave and Bororo the terms in question are not so analyzable. Two languages of South America, Arabela and Yanomámi, colexify 'flame' and 'candle' (and Arabela colexifies also 'match' and 'wax'). Relatedly, Wintu *sayi* also colexifies 'lantern' and 'lamp,' and Gurindji employs the same term for 'flame' and 'torch' (see Buck 1949: 72 for parallel evidence from the history of Romance). The Buli, Badaga, Aymara, Cashinahua, Hupda, and Piro terms also include 'embers' in their denotational range. Khoekhoe *llhabub* ~ *llkhabub* is derived from *llhabu* 'to catch alight,' and 'flame, flare up.' The latter meaning is also colexified in Muna, and there is a semianalyzable term in Rotokas. Lesser Antillean Creole French, Pawnee, and Hani colexify 'blaze' (Hani also 'ardor'). In Hawaiian and Kapingamarangi, relevant terms also can refer to species of lobsters.

Other associations include: Dongolese Nubain colexifies 'flame' with 'flower, blossom.' The Kyaka term for 'flame' varies in form between *uu lenge*, *wii lenge* and *wilenge*. All of these terms apparently contain the Kyaka word for 'eye,' *lenge* (which, however, also means 'node, knuckle' inter alia). Sahu *lejanga* can also refer to a 'glow' inter alia. Badaga *kiccu* is also used figuratively with the meaning 'jealousy,' and *kolli* is glossed as "'thicket, bush, firewood; flame, brand, firebrand,'" Basque *sugar* may also refer to a 'male snake,' and *lama* also means 'glare, gleam' alongside -accidentally- 'lama,' a spiritual teacher in Buddhism. Greek *flóga* also has the figurative meaning 'fire, passion,' Khalkha *zali* can also mean 'spirit' and refer to "[r]use, craft, cunning, trick, deceit," and *doly* can also mean 'to be timid, not dare.' Itzaj *jom* also means 'to break hole, perforate,' Nez Perce *?á'la* also 'hell,' and Wintu *čul* may be related to *čul* 'pour, spill.' The Chayahuita term *ohuica pochin nininso* is apparently based on a metaphorical connection with sheep (*ohoica*), which flames are apparently felt to be (*ninin*) like (*pochin*), and for Embera *ne eráadriú*, compare *eraadrú* 'lightning.' 'Flame' and 'lightning' are indeed colexified in Tsafiki. Guaraní *tendy* also means 'saliva.' Toba *l'edaxa* colexifies 'brilliance, gleaming' alongside 'heat, fever,' and similarly, Tehuelche colexifies flame with 'flickering.' Tsafiki *pinda* also means 'thunderbolt.' Hani *miqlaol* contains *laol*, a verb meaning 'be warm or hot,' and *miqbia* contains *bia* 'bright, shining, flash.' Great Andamanese *archâl* is derived from *châl* 'beam, shine.'

23. *The Flood*

Representation: 49%

Motivated: 49.6%

Thereof Analyzable: 30.6%

Thereof Colexifying: 18.5%

Thereof by Contiguity: 28.1%

Thereof by Similarity: 10.2%

Recurrent associated meanings: water, river/stream, torrent, overflow, come/move/go inland, big, current, flow, push underwater, cover, sea, swell, (be) full, scrape

Terms for ‘flood’ (or ‘floodwater,’ ‘deluge,’ ‘flood tide’) betray a variety of lexico-semantic associations cross-linguistically. Frequently, the meanings ‘water’ or ‘sea’ act as contiguity anchors in morphologically complex terms, and in four further languages, Berik, Yir Yoront (where an etymological connection with ‘sand’ may be present), Comanche, and Piro, semianalyzable terms are found where one of the constituents with the meaning ‘water’ is identifiable.

In five sampled languages, Kyaka, One, Upper Chehalis (in One and Upper Chehalis there is a single term for ‘water’ and ‘river,’ and hence, there is also a lexical association with ‘river;’ similarly, Ngaanyatjarra and Badaga colexify ‘floodwater’ with ‘river’ or ‘stream,’ and Ineseño Chumash ‘stream, creek’ with ‘flood’ itself), Chickasaw and Lake Miwok, the second element in their complex terms for ‘flood’ is ‘big’ or ‘to be big.’ For instance, Lake Miwok has *ʔudíkik* /ʔudí-kik/ ‘be.huge-water.’ Note also Efik *a’qua i’nyañ* ‘great wide.expanse.of.water’ and the formally redundant Central Yup’ik term *ule-rpak* ‘flood/high.tide-big.’ In five languages, Dadibi, Sko, Nivkh, Pawnee, and Hawaiian, the second term has the meaning of ‘come,’ ‘go inland’ or ‘move forward’ (for instance Dadibi *asobo wé* ‘come water,’ and Sko *tí hoe toe* /tí hóe toe/ ‘sea go.beachwards come.up’). In the Pawnee term *kicuuta’a*, an additional element with the meaning ‘swell’ is encountered; the term is analyzable as /kic-huuta’at/ ‘be.liquid-swell’ (compare colexification of ‘to flood’ and ‘to swell’ in Jarawara). Similarly, Haida has *gay-hll* ‘ACTION.OF.WAVES-move’ (there is another term featuring *gay*). In Yoruba, there are the terms *ì-kún omi* ‘NMLZ-be.full water’ and *ì-sàn omi* ‘NMLZ-flow water.’ The former association with ‘full’ is mirrored in Rotuman by colexification (where the relevant term also has many other meanings), and the latter is also encountered by derivation in Samoan, in Khoekhoe, where *dāus*, a term for the Biblical flood, is derived from the root *dāu*, meaning ‘flow, stream, run,’ and furthermore, in Nuuchahnulth, where *cuupšĩł* is analyzable as /cũp-šĩł/ ‘for.liquid.to.flow.out-MOMENTANEOUS.’ In this context, note also the similarity between Rendille *dúley* ‘flood’ and *dula*, meaning inter alia ‘flow, flood, be flooded,’ as well as that between Yuki *ʔal-* ‘flood’ and *ta-* ‘to flow.’ Other analyzable terms with ‘water’ or ‘sea’ as contiguity anchors include Koyraboro Senni *harihurey* ‘high flood (in annual cycle),’ which contains *hari* ‘water, liquid’ and *huru* ‘to enter,’ Baruya *aalya-aka* analyzable as ‘water-white’ and colexifying ‘muddy water,’ Ket *ákul* (analyzable as /aq-ũl/ ‘rot-water’ and also meaning ‘hole in ice’), and Cheyenne *māhóovátó*, which is said to mean ‘all over water’ literally. Kashaya has *ʔahq<sup>h</sup>a co’bi?* (/ʔahq<sup>h</sup>a <sup>h</sup>co-Xi<sup>ʰ</sup>bi<sup>ʰ</sup>c/ ‘water carry-up’), Guaraní *y-sẽ* ‘water-leave,’ and Hawaiian *wai-holomoku* ‘water-rush.’ Somewhat unclear and potentially spurious associations with ‘sea’ acting as contiguity anchor include San Mateo del Mar Huave *tenguial ndec*, where

*ndec* is 'sea' and *tenguial* is a demonstrative element and Rama *táuli alka bángi* 'salt/sea sun/hot let.us' (One also has a term colexifying 'floodwaters' with 'sea,' and there is a semianalyzable term where the identifiable constituent means 'sea' in Miskito). In addition, as mentioned above, three languages in the sample, Ngaanyatjarra, Badaga, and Ineseño Chumash, colexify 'flood' or 'floodwater' and 'river, stream' (Badaga also 'ditch,' 'lake,' and 'bottomland' inter alia), while in Tetun, one term for 'flood' is *mota-tun* 'river-down/descend,' and there is a semianalyzable term where the identifiable constituent means 'water, river, lake' in Berik and one where it means 'river' only in Rama. In two languages of New Guinea, Kwoma and Kosarek Yale, there are verbs colexifying the meanings 'to cover' and 'to flood' (Kosarek Yale also colexifying 'to hide' and 'to keep secret'), and in Tuscarora, the term for 'flood' consists of the contiguity anchor 'water' and a stative verb 'cover' with the term for 'land' incorporated (*à'wə? wa?wnawérhə*, analyzable as */à'wə? w-a?wT-awerhu-ə/* 'water 3SG.NEUT.AGENT-land-cover-STAT'). Also in Kwoma and Kosarek Yale, there are verbs meaning 'to push underwater,' 'to put under surface of water' and 'to flood;' Bislama has *draonem* (< Engl. *drown*) with the same semantic structure, and, given the language's relative proximity, it is intriguing to speculate whether this is a more general semantic pattern of the broader New Guinea area which Bislama has acquired by relexification. Similarly, 'flood' and '(to) torrent' are colexified in Kwoma, Kyaka, Basque, Guaraní, Hawaiian, and Tetun, and 'flood' and '(to) overflow' are in Buli, Muna, Comanche, Nuuchahnulth, and Hawaiian. In two languages of Oceania, Mali and Fijian, complex terms for 'flood' revolve around a verbal element meaning 'to scrape' (see Stebbins n.d. for details of the underlying conceptualization in the former). In Buin, Kwoma, and Welsh, associations with the meaning 'current' are encountered (the relevant Welsh term *llif* also means 'saw'). In Kwoma, *ukwi kiya* is analyzable as 'water.current carry,' and in the other languages the association is by colexification (Buin also colexifies "be light brown (the colour of a river in flood)").

Other associations include: Buli *mobi*, used verbally, means 'to cut,' 'to crack,' and 'to burst' (said of banks) inter alia, Hausa colexifies "ingress of mass of water" with "watersprout, cloudburst" as well as 'congestion.' Kwoma has two colexifying terms, namely *dabu*, which also means 'to pour down' or 'to spill' inter alia, and the aforementioned *ukwi kiya*, which means 'for a river to rise' and again 'to pour down.' Meyah colexifies 'flood' with 'surplus' and Kosarek Yale *yamak-* also means "to fill the air with a smell." Basque *uholde* can also be used metaphorically to a plethora of something (as can English *flood*). The relevant Chukchi term is semianalyzable, containing an element meaning 'deep,' and Yana *zuu-* also means 'to push' and 'to poke, to spear.' Bislama *draonem* also means 'to post a letter.' Hani *eelpuv puv* contains *puv* meaning inter alia 'to roll, to cross over,' *eeltaoq taoq* contains *taoq* 'to pound, to butt, to rub against,' and *eelpuv leiq* also contains *leiq* meaning 'to look for, to deviate' inter alia. Kapingamarangi *doloo* also denotes a species of duck. Samoan *lolo* is a nominalization of a verb meaning 'for water to run,' and Tetun *nabeen* can either mean 'to liquify' or 'to flood.'

24. *The Foam*

Representation: 85%

Motivated: 32.5%

Thereof Analyzable: 11.5%

Thereof Colexifying: 21%

Thereof by Contiguity: 13.6%

Thereof by Similarity: 13.9%

Recurrent associated meanings: bubbles, saliva/spittle, suds, water, lungs, soap, boil, spray, bud, fog, foam on mouth, swell

The meaning ‘foam’ (‘froth,’ ‘scum’), when terms for it are lexically motivated, is expressed to about equal proportion by terms with an underlying semantic relation of contiguity and of similarity, and this is reflected in the two cross-linguistically most frequently associated meanings: ‘bubbles’ by contiguity and ‘saliva, spittle’ by similarity. The association with ‘bubbles’ occurs in fourteen languages by colexification, Mbum, Gurindji (colexifying also ‘steam’), Kwoma, Nunggubuyu, Ineseño Chumash, Haida, Itzaj, Cayapa, Embera, Guaraní, Wayampi, Hawaiian, Mandarin, and Rotuman, and one language, White Hmong, has the analyzable term *npuas dej* ‘bubble water,’ which, however, denotes ‘bubbles’ itself as well. Noteworthy is the Guaraní term *tyjúi*, perhaps analyzable as /ty-jýi/ ‘urine-rainbow,’ which then would make reference to the iridescent reflections of the light occurring especially with bubbles on soapy water (in fact, an association with ‘soap’ or ‘soap powder’ is found in Ngaanyatjarra, Sko, and Miskito, which latter has the analyzable term *sōp laya* ‘soap water,’ and ‘suds’ is colexified in Koyraboro Senni, Burarra, Khalkha, Cahuilla, Pawnee, Aymara, Hawaiian, and Rotuman, while Nez Perce *tí’píp* is related to *tipí’pi* ‘become sudsy’).

Interesting is the association with ‘saliva, spittle’ or sometimes ‘slave, drivel’ because it is particularly frequent in New Guinea and Oceania. It is found in a total of thirteen languages, Baruya, Buin, Kwoma, Lavukaleve, Muna, Ngaanyatjarra, Sko, Southeastern and Western Tasmanian, Badaga, Lesser Antillean Creole French, Tsafiki, Bislama, Lenakel, by colexification in nine languages and by analyzable terms in four (Baruya has *aali-maagwala* ‘water-saliva,’ Muna *bura-no tehi* ‘foam/froth-poss sea,’ where the example in the source suggests that *bura* is used for ‘foam on mouth,’ but this case remains somewhat unclear, Sko *fōefōe*, reduplicated from *fōe* ‘spittle,’ and Lenakel *nihi-noua tehe* ‘liquid-mouth sea,’ where *nini-noua* ‘liquid-mouth-’ is ‘saliva’). Similarly, Basque and Malagasy colexify ‘foam on mouth’ and Kaluli “froth from mouth during seizure” specifically. Given that ‘saliva’ has a foamy structure as well, and is also in a contiguous relationship with ‘foam at the mouth’ one may wonder whether in the colexifying languages, the target meaning is really ‘foam’ as found on water. Indeed, morphologically complex terms like that found in Lenakel suggest that this may be so (although, to be sure, each language is different and there may be some spurious cases).

In Buin, Kaluli, and Toaripi, ‘foam’ is colexified with ‘lungs’ (due to the spongy appearance of this organ), and Kaingang *jēngéj* is glossed as ‘foam of lung’ (it is also common cross-linguistically to have complex terms for ‘lungs’ on the basis of ‘foam,’ see section 122). Two languages, Khoekhoe and Ancash Quechua, colexify the meanings ‘foam’ and ‘bud’ (in Khoekhoe, especially the bud of *Acacia Watkins*), and another two, Cahuilla



and Hawaiian, colexify 'foam' and 'fog.' Similarly, Ngambay colexifies 'fog' and Gurindji 'steam.' Also in two languages, an association with a verb meaning 'to boil' is found. In Itzaj, the same term may be used as a verb with the meaning 'to boil' and as a noun with the meaning 'foam,' and in Abipón, *l-apa-Ra* 'steam of boiling water, foam' is analyzable as 'POSS.INDEF/3SG-boil-ABSTR.' Efik *ěfut* (probably accidentally also meaning 'fifteen, fifteenth') is derived from a verb colexifying 'to boil' with 'to swell, ferment, foam' and other meanings, and similarly, Khalkha *køgesy(n)* is a resultative nominalization of the verb *køge* 'to swell, distend, intumescence, foam.' In addition, the Xicotepec de Juárez Totonac term *puput* 'foam' appears to be derived from *pupú* 'to boil,' and lexical connections between 'foam' and 'to boil' are also reported for Wintu and, diachronically, for Haida. Furthermore, Burarra and Wayampi colexify 'foam' with 'spray' on waves, and in Great Andamanese, the same root *bôag* yields the meaning 'foam on mouth' and 'foam on sea' depending on the possessive prefix attached.

Other associations include: Buli *puuk* also means 'stomach, belly' and 'pregnancy' inter alia, Hausa *kumfa* also 'and then, thereupon,' and Burarra *munjanachana* is derived from *janachana* 'make crumbs' by prefixation of the class marker *mun-*. Rotokas *ruiruiso* appears to contain *ruirui*, meaning 'to dry up' or 'to wipe up,' Nunggubuyu colexifies "flood debris," and Toaripi *ma sese* 'slight foam on waves, white caps' is analyzable as 'water fibres' (the Northern dialect of Ngaanyatjarra colexifies 'foam' with 'water' directly, there is a semianalyzable term with an element meaning 'water, river' in One, and a Ket term may be diachronically relatable to one meaning 'water' as well). Abzakh Adyghe *tx°ambe* (with the variant *tx°arbe*) is analyzable as /tx°ə-m-be/ 'white-RELAT/EPEN-hollow.' Greek is the only language in the sample where the term for 'foam,' *afrós*, can also mean 'cream,' Japanese *awa* also means 'foxtail millet,' Khalkha *ceger* also 'taboo, prohibition, abstinence' and 'quarantine,' and Sora *bub'bu'da:n* also "[b]eads of various kinds made of glass or metal, used as ornaments." Kolyma Yukaghir colexifies "dirty water which remains after cleaning a fish." Itzaj *otz'tik* and *ootz'* can also mean 'wrinkle,' Lake Miwok *póta* also 'to be gray, to be cloudy' and 'semen,' Tuscarora colexifies 'effervescence' with 'thin foam,' Cavineña colexifies 'tree,' and Cayapa *chimbijpu* (or *shimbijpu*) also denotes the 'bladder' of an animal. Ancash Quechua *ñawi* also means 'eye,' 'spring of water,' and has other related readings. Rama *ngú aya* and *ingu aya* literally mean 'house corn.' The Toba term for 'foam' varies in form between *lchi* and *lcochi*; *lchi* also means 'leg, tributary.' Fijian *vuso* generally also means 'top end' and 'to squeeze out the juice from leaves, coffee powder.' Hawaiian *hu'a* is, presumably by metaphor, extended to 'border' and 'suburb' among other meanings, and *ehu* also means 'dust,' 'pollen,' and 'faint, difficult to see' inter alia. The complex Manange term *2kju 1atsaŋpa 1mo 1mu* contains *2kju* 'water,' *tsaŋ* 'clean' and a negative marker inter alia, and Mandarin colexifies 'foam, bubble' with 'blister' as well as 'to soak.' Tetun has *furi-n* 'sprinkle-SINGULATIVE,' and Samoan *piapia* is reduplicated from *pia* 'secretion of the genital organs, smegma.' Bislama *nus* (< Engl. *nose*) means 'nose,' 'mucus of nose' (by metonymy) and 'foam' (by metaphor).

25. *The Fog*

Representation: 89%

Motivated: 47.7%

Thereof Analyzable: 13.4%

Thereof Colexifying: 34.5%

Thereof by Contiguity: 1.6%

Thereof by Similarity: 35.7%

Recurrent associated meanings: cloud, smoke, steam, dew, dust, cover, white, cold/frost, darkness/darkening, foam, rain, fall, land/ground, water

For the meaning ‘fog,’ associations with other aerosols, namely ‘smoke,’ ‘steam,’ and ‘cloud’ abound. These associations are frequently realized formally by colexification, but also quite often by morphologically complex terms. The interesting fact, however, is that the relationship between the meanings as revealed by analyzable items is unidirectional, in that there are complex terms for ‘fog’ on the basis of the aforementioned meanings, but the reverse situation is unattested in the sample.

The association with ‘cloud’ is present by colexification in Bakueri, Buin, Ngambay, Dongolese Nubian, Yoruba, Anggor, Baruya, Kwoma, Meyah, Ngaanyatjarra, Rotokas, Waris, Kosarek Yale, Basque, Bezhta, Sora, Haida, Lesser Antillean Creole French, Nez Perce (by a lexical affix), Xicotepec de Juárez Totonac, Arabela, Aymara, Maxakalí, Piro, Ancash and Imbabura Quechua, Tehuelche, Yanomámi, Manange, Sedang, Takia (and perhaps Tasmanian, Plomley 1976: 230, as well as perhaps Embera, where the relevant terms are almost identical segmentally). In Ngaanyatjarra, Sedang, and Takia, the colexified meaning is ‘low cloud’ more specifically, while even more specialized types of cloud are colexified in Buin (‘white cloud’), Aymara (‘fairly dark cloud’), Chayahuita (‘white cloud’; there is also an analyzable term of the redundant type with the additional constituent meaning ‘to rise up’), Hani (‘thick cloud’), and Hawaiian (‘light cloud on mountain’). Similarly, Badaga colexifies ‘veil of cloud.’

Analyzable terms include Kyaka *yangama kopa* and *yuu kupa* ‘morning cloud’ and ‘ground cloud’ (both denoting ‘ground fog’ specifically), Upper Chehalis *s-pát=šq* ‘CONTINUATIVE-drop/fall=cloud,’ Highland Chontal *dedíhima lummaway* ‘encircle cloud,’ Santiago Mexquititlan Otomí *’bongui* /’bóni-gui/ ‘be.stretched.out-cloud,’ White Hmong *pos huab* ‘moist cloud,’ and Samoan *pu-ao* ‘hole-cloud’ (see Buck 1949: 66 for the association between ‘fog’ and ‘cloud’ in Indo-European, which is common there). In Khoekhoe, there is a semianalyzable term where the identifiable constituent is ‘stratus cloud.’

‘Fog’ and ‘smoke’ are colexified in Efik (by the analyzable term *nsuñ’ikañ* /n-suñ-ikañ/ ‘soft/gentle-fire’), Buin (‘white smoke’ specifically is colexified here), Burarra, Gurindji (colexifying ‘light fog’ with ‘smoke haze’), Yir Yoront, Abzakh Adyghe, Nez Perce (by a lexical affix), Wintu, Central Yup’ik (colexifying also ‘dust in air’), Hupda, Jarawara, Lengua, Maxakalí, Miskito, Wayampi (by the analyzable term *atāsī* /ata-sī/ ‘fire-whiteness’), and Mandarin (colexifying also ‘cigarette, tobacco’). There are semianalyzable terms in Ngaanyatjarra (here, the term also denotes a ‘grey-leaved acacia shrub’), One, San Lucas Quiaviní Zapotec, Lengua, and Rama. Similarly, Burarra has another term colexifying ‘light fog’ with ‘smoke screen.’ Analyzable terms are Santiago Mexquititlan Otomí

'bipa, analyzable as /'bifi-pa/ 'smoke-heat,' and Arabela *cohuaja najaca* as 'white cloud/smoke/vapor.'

Finally, 'steam' (or 'vapor' generally) is colexified in Efik (again by the analyzable term *nsuñ'ikañ* /n-suñ-ikañ/ 'soft/gentle-fire'), Anggor, Burarra, Kyaka (also colexifying 'pawpaw'), Yir Yoront, Abzakh Adyghe, Greek, Khalkha, Welsh, Wintu (colexifying also 'gas' and 'lungs'), Arabela, Hupda, Lengua, Miskito, Hawaiian, and Rotuman, which also colexifies 'fine spray' (see also Buck 1949: 67 for evidence from Swedish and Norwegian). Perhaps similarly, Sora colexifies 'black vapour' more specifically. There are two languages with analyzable terms: Ket has *ulij* /ul-ii/ 'water-vapor' and Lesser Antillean Creole French *vape ki ka kouve late ében lanme*, containing *vape* 'steam,' *kouve* 'cover,' and *lanme* 'sea.' As discussed in some more detail in § 6.2.2.2., in the colexifying languages, sometimes more than one of the associated meanings is expressed by the same term.

Furthermore, a lexical association between 'fog' and 'dew' is found in eleven languages of the sample (see also Buck 1949: 66 for evidence from Greek): Burarra, Rotokas, Biloxi, San Mateo del Mar Huave, Quileute, Wintu, Arabela, Rama, Hawaiian (where the relevant term is also the name of a "gentle land breeze, as of West Hawai'i"), Bwe Karen and Vietnamese. The association is mostly by colexification (in Burarra, the meaning colexified is more precisely 'heavy fog,' and in Wintu, 'dewy' is colexified with 'foggy'), with the exception of Vietnamese, where 'fog' is *sương mù* 'dew blind' and Biloxi, which has *ayuxka'*, presumably /ayu-ka/ 'dew-something.' The reasons for this association remain somewhat unclear, although a likely source for it is that both 'fog' and 'dew' require low temperatures to occur, which would then be a case of motivation by temporal contiguity. Some circumstantial support comes from the San Mateo del Mar Huave term *ajmiüc naquind*, which, in addition to colexifying 'fog,' is also analyzable as 'fall cold' (compare the Upper Chehalis term involving a constituent meaning 'to drop, fall'). In fact, in two further sampled languages, Yuki and Bislama, the meaning 'fog' is expressed by morphologically complex terms associating it with 'cold' or 'frost' (which is the cause for the phenomenon in the first place). Yuki *kō nq̣t* contains or is related to *kōh* 'frost' and *nq̣t* 'ice,' while Bislama *kolkol* (the reduplication base being *kol* 'cold') can itself also mean 'cold,' 'cool,' 'stale' (said of food) and other things as well, and 'be foggy' is moreover colexified with 'be frosty' in Blackfoot. An association between the color 'white' and the meaning 'fog' is found in four sampled languages, namely Kwoma (here by colexification) as well as Chickasaw, Arabela, and Wayampi. Chickasaw *tooboklhili* ~ *tohboklhili* is analyzable as /tohbi-oklhili/ 'be.white/be.pale-night,' Arabela *cohuaja najaca* as 'white cloud/smoke/vapour,' and Wayampi *atāsi*, as already mentioned above, as /ata-sĩ/ 'fire-whiteness' (the Wayampi term, as suggested by the semantics of its constituents, also colexifies 'smoke,' so it is dubious whether the association of 'white' with 'fog' should be considered as genuine in this case). Note also Rendille *'duubāt* 'fog' and *dúub* "white circular hat, turban of white cloth."

Complex terms for the meaning which involve a verbal element with the meaning 'to cover' are found in four sampled languages, Cheyenne, Kiliwa, Lesser Antillean Creole French, and Hawaiian. However, the conceptualization seems to differ to some extent in spite of this commonality: in Cheyenne, *nēhpóemáno'e* can be literally translated according

to lexicographers as “closed.in(covered)-environment,” whereas in Kiliwa, *yuw=hi?* is analyzable as ‘eye=cover.’ Hawaiian has *uhi-wai* ‘covering-water’ (colexifying also a “kind of tapa”); for the Lesser Antillean Creole French evidence see above. In Kolyma Yukaghir, in addition, *šazil* ~ *šaril* can refer both to ‘fog’ and a ‘covering,’ and indeed is derived from the root *šar-* “to cover, to bury, to press; to overtake,” and can, presumably due to the sense ‘bury,’ also refer to a ‘root.’ Two languages of Eurasia, Khalkha and Welsh, colexify ‘fog’ and ‘darkness,’ ‘dimness’ or ‘darkening,’ featuring at least one term (alongside others) that can express both meanings. Similarly, Nez Perce has a word for ‘fog’ as well as “darkening from fog or approaching storm” which contains the lexical affix *ʔipé-* ‘pertaining to smoke, clouds, fog’ and is hence formally redundant, but contains a further constituent meaning ‘to be dark.’ Wayampi *tewilaki* contains *tewi* ‘buttocks, rump’ and *ki* ‘to rain,’ Dadibi colexifies ‘mist’ and ‘slight rain,’ and a semianalyzable term one of the constituents of which is ‘rain’ is also found in Noni. Relatedly, associations with ‘water’ acting as a contiguity anchor are found in Ket and Hawaiian, as mentioned already, and there is a semianalyzable term in Itzaj. In two languages of the sample, Cahuilla and Hawaiian, ‘fog, mist’ and ‘foam’ are colexified (Cahuilla also has another term related to a verb meaning ‘to enter’). In four languages, Abzakh Adyghe, Wintu, Hupda, and Hawaiian ‘dust(y)’ and ‘fog(gy)’ are colexified (although it should be noted that the semantic range of the Hupda term is quite large, also colexifying ‘smoke’ and ‘steam’), and Itzaj has *aj-tanat-ja* ‘MASC-dust-water.’ Similarly, Central Yup’ik colexifies ‘dust in air’ specifically with ‘haze.’

Other patterns include: Dongolese Nubian *níčč(i)* is related to *nĩǵ* ‘to sew’ and also denotes the ‘action of sewing;’ the extension to ‘cloud, mist’ is explainable according to the source by clouds and fog “veiling” the environment like what one sews, that is, cloth. The Rotokas terms *guiguisiva* and *vusiva* may be related to the verbs *guiguisi* ‘spray out’ and *vusi* ‘burst forth’ respectively, compare the apparent lexical relationship between Dadibi *segeni ebo* ‘mist, slight rain’ and *sege ebo* ‘heavy.’ Muna *gawu*, as a verb, also means ‘be hazy, dim, distant’ said of objects or vision, Western Tasmanian perhaps colexifies ‘fog’ with ‘shadow,’ while Yir Yoront also colexifies ‘spray on waves,’ and similarly, Hawaiian ‘rain spray.’ Abzakh Adyghe *-y°e-* also means ‘yellow, blond’ inter alia, and Basque *laino* is also used metaphorically with the meaning ‘confusion.’ Japanese *kiri* also means ‘awl’ (though *kiri* ‘fog’ is a nominalization of *kir* ‘become foggy,’ so the similarity is accidental). A precise morphological analysis of Cheyenne *nêhpoése ma’ěno* is not possible; however the literal meaning of this term according to the lexicographers is ‘the turtle is shrouded.’ Guaraní *tatatîna* contains *tata* ‘fire.’ Ancash Quechua colexifies ‘moth,’ Fijian *kabu* also means ‘to sow or scatter small seeds,’ and for Bwe Karen (*de*)*θochí*, perhaps compare *θochí* ‘gallbladder.’ In Bislama, the meaning of *sno* (< Engl. *snow*) has been extended to also mean ‘fog.’ Hawaiian *ohu* colexifies “adorned with a leis,” and *ehu* inter alia also ‘faint, difficult to see’ and ‘pollen.’ Finally, Tetun has *rai-ahu* ‘land-lime.’

26. *The Forest*

Representation: 84%

Motivated: 42.1%

Thereof Analyzable: 10.9%

Thereof Colexifying: 31.3%

Thereof by Contiguity: 21.2%

Thereof by Similarity: 0%

Recurrent associated meanings: bush/grove/scrub, wood, tree, mountain/hill, place/area, lumber/timber, branch/stick, plant, weeds, wilderness, thicket, taiga, place to defecate, grass, interior

As one might expect, terms for 'forest' (or 'jungle') are most frequently lexically associated with terms for 'tree' and/or 'wood,' either by colexification or morphological (semi-)analyzability (see Buck 1949 for evidence from Indo-European, among the languages where this association is attested diachronically is German). For 'tree' specifically, colexification is found in six languages: Ngambay, Khalkha, Upper Chehalis, Nez Perce, Yaqui, and Hawaiian (Ngambay, Yaqui, and Hawaiian, like Muna and Wayampi, colexify also 'plant,' and the Hawaiian term has also still other meanings). Alternatively, seemingly analyzable terms for 'forest' on the basis of 'tree' are found in eleven sampled languages, Anggor, Baruya, Kwoma, Comanche, San Mateo del Mar Huave, Santiago Mexquititlan Otomí, Arabela, Ancash Quechua, Fijian, in another Hawaiian term for the meaning, as well as in Tetun. In Baruya and Ancash Quechua, the relevant terms are of the lexical type, the second element bearing the general meaning of 'place' or 'area' (Baruya *yí'darya*, literally 'tree area' for 'clump of trees, forest' and Ancash Quechua *sacha marka* 'tree/plant area'), while in Anggor, *nimambe* 'forest, jungle' is analyzable as /nimi-ambe/ 'tree-in,' Kwoma *me kaba* "mature forest (including sago swamps) growing on relatively level terrain" might be a dvandva compound of *me* 'tree, wood' and *kaba* "type of long-bladed grass that grows prolifically in swamps and lagoons," in Comanche, the word for forest, *soo huuhpi*, is analyzable as 'many trees,' in San Mateo del Mar Huave, *tixiül* appears to be analyzable as 'AUG-tree,' Pipil (Cuisnahuat dialect) has *ku(:)htan*, analyzable as /kuh-tan/ 'tree/wood-locative/under,' and in Santiago Mexquititlan Otomí, 'forest' is *mbo-zaa* 'place.where.around-tree.' The Arabela term is formally the plural of that for 'tree,' Fijian has *vei-kau* 'COLL-tree/wood/stick,' and Tetun *ai-laran* 'tree-interior' (compare colexification of 'jungle' and 'interior' in Berik, presumably because this is where it is located). Hawaiian, moreover, has *ulu lā'au* 'grove tree/forest' (there is a further rare semianalyzable term featuring *ulu* which also means 'brush, undergrowth'), and there is another term that can mean either 'trees' or 'forest' (note that for Arabela and Cashinahua, the respective terms are also glossed as 'trees'). A Sora term for 'forest' is *ə'ra:ʃa:ŋən*, containing *ə'ra:* 'wood, wooden' and perhaps *ʃa:ŋ-* 'bone, hard dry stalk.' In addition, there is a relatively large number of languages with semianalyzable terms for the meaning 'forest,' one of the constituents clearly being words for 'tree,' namely Kaluli, Xicotepec de Juárez Totonac, Bororo, Embera, and Sáliba.

Colexification with 'wood' is found in fifteen languages in the sample: Buli, Ngambay, Dongolese Nubian, Badaga, Basque, Chukchi, Greek, Khalkha, Laz, Welsh, Haida, Miskito, Hawaiian, Tetun, as well as Lesser Antillean Creole French (where the relevant

term also means ‘arm’). This situation may also obtain in Wayampi, though here the meta-language gloss ‘bois’ itself colexifies ‘forest’ and ‘wood,’ thus making the case unclear. Cubeo has *jocu-burumu* ‘wood-vegetation.that.is.not.cut.down,’ Fijian, as mentioned above, *vei-kau* ‘COLL-tree/wood/stick,’ and there are semianalyzable terms with an identifiable constituent meaning ‘wood’ in Sora, Xicotepec de Juárez Totonac and Maxakalí. Similarly, Khalkha, Cheyenne, Nez Perce, Pawnee, Hawaiian, and Lesser Antillean Creole French (also) colexify ‘lumber, timber’ more specifically.

As also discussed in § 6.3. dealing with environmental factors and their influence on patterns of colexification, eight languages colexify ‘forest’ and ‘mountain’ or ‘hill’ (see also Buck 1949: 46-47 for evidence from Indo-European): these are Kwoma, Laz, Nivkh, Sora (also colexifying ‘clearing on hill’), Ineseño Chumash, Aguaruna, Huambisa, and Yay. Moreover, Kwoma colexifies ‘forest’ also with ‘netbag’ and ‘womb,’ Basque also with ‘wooded hill,’ and Khalkha also with “mountain ridge, range, or plateau,” alongside ‘taiga,’ the latter association being shared with Kolyma Yukaghir (note also that the potential cases of Xicotepec de Juárez Totonac, Copainalá Zoque, Aguaruna, Arabela, Aymara, Bora, Cashinahua, Cavineña, Cayapa, Chayahuita, Cubeo, Guaraní, Huambisa, Lengua, Sáliba, Tsafiki, and Yanomámi, where the relevant terms are (also) glossed as ‘monte’ and/or ‘selva,’ which are capable of referring to a ‘mountain’ but more generally denote all kinds of non-cultivated territory, and are hence ignored here when assessing colexification).

Quite a number of languages colexify ‘forest’ with meanings such as ‘bush, bushland,’ ‘grove,’ or ‘scrub, undergrowth, underbrush.’ These are: Bakueri, Buli (also colexifying ‘savannah’), Efik, Hausa, Ngambay, Buin, Burarra, Gurindji, Kwoma, Nunggubuyu, Sko, Northeastern Tasmanian, Badaga, Blackfoot, Cheyenne, Haida, Ineseño Chumash (also colexifying ‘chaparral,’ a type of vegetation in California), Nez Perce, Pawnee (by a semianalyzable term containing an element meaning ‘place’), Chayahuita, Miskito (also colexifying ‘opening, hole’ as well as ‘rectum, anus’), Rama, Wichí, Hawaiian, Lenakel, and Samoan; Kanuri *kàràà sálám* is literally ‘bush black.’ Moreover, Oneida *kélhite?* is analyzable as /ke-lh-Nít-e?/ ‘ANOMALOUS.PREFIX-woods-be.in-STAT.’ Chickasaw *abokkolanonka* is analyzable as /abokkoli'-anonka/ ‘thicket-inside,’ while Khalkha and Hawaiian colexify ‘forest’ and ‘thicket’ directly.

Three Austronesian languages, Muna, Hawaiian, and Samoan, colexify ‘forest’ and ‘weeds,’ Muna and Wayampi also ‘grass,’ and Samoan ‘tall grass’ specifically. Furthermore, Ngambay colexifies ‘branch,’ and Khalkha, Nez Perce, Bororo, and Hawaiian colexify ‘forest’ with ‘stick’ (Nez Perce with ‘small forest’ more specifically), while in Carrier, *tcentherh* is analyzable as /tetcen-therh/ ‘stick/wood-in.’ While for Ngambay, Khalkha, and Nez Perce, it should be noted that the relevant terms also colexify ‘tree’ and the association with ‘branch’ may thus be not immediate, this is not the case for Bororo. Ngambay *kake* in fact has a very broad semantic range and may also refer to a ‘rod,’ ‘stick,’ as well as ‘place’ generally, a pattern shared with Jarawara, where the relevant term also means ‘thing’ and ‘time’ (see § 6.4.3.15.), and Kaingang. Similarly, Yanomámi colexifies ‘forest’ with ‘region, area, territory’ as well as ‘city’ and ‘country,’ and note also that the semianalyzable Lengua term mentioned above contains an element meaning ‘place.’ In Badaga, *ka:du*, meaning ‘forest, jungle woods,’ but also ‘field’ and ‘wasteland’ among other meanings, is also used

as a term for the “place where one defecates,” and in Lenakel, *nikinhamra* ‘forest, bush’ is also glossed as “[t]he place in the bush where one goes to urinate or defecate ...” Finally, relevant terms in Koyraboro Senni, Badaga, and Hawaiian are also glossed as ‘wilderness.’

Other associations for this meaning and semantic nuances conveyed by the respective terms are manifold. Buli *sagi* also can refer to a far-away country inter alia, Hausa *kurmi* also denotes a “wooded ravine” and is an “expression of wonderment,” and Ngambay *kər* also means ‘coal’ inter alia. Buin *maika* also is the name of two Buin villages, the Burarra term *gumurra* is based on the root *murra* meaning ‘clustered.’ Kwoma colexifies “dense undergrowth” among other meanings. The Kyaka term *imwua* can also refer to the ridge-pole of a house that is cut from a very tall tree, and Muna has *katugha* ~ *kamothuga*, derived from *tugha* ‘hard, ripe’ and *karumbu* ~ *karombu* ~ *kahombu*, perhaps related to *rumbu* ‘to stink (as of old wounds).’ Another term colexifies ‘forest, jungle’ with “part of a field not yet cleared (of weeds),” as well as to “separate off part (of a field) to be worked later.” Sko *hángpeng* apparently contains *peng*, which can either refer to a tree species or mean ‘clear, cleared,’ and Yir Yoront *larrkurraq* contains *larr* which means ‘land, site’ and other related things. Basque *baso* also means ‘vase, glass’ and ‘vessel,’ these senses are presumably due to borrowing from Romance. The Khalkha term *cabdayul* is derived from *cabda* ‘split,’ using the agent nominalizer *-yul* and also means ‘grove’ as well as ‘domino,’ while *oi* also means ‘birthday, anniversary,’ and ‘mind, intellect, memory’ inter alia. Upper Chehalis *yámacumš* appears to be analyzable as /*yámac*=*umš*/ ‘douglas.fir=place,’ while Nez Perce colexifies ‘forest’ with ‘tree trunk.’ The Wintu term *kel* can also be used in an adjectival sense, then meaning ‘long, far.’ Highland Chontal colexifies ‘forest’ with “wide open space outside the village, the woods,” and more specifically also ‘airstrip’ and ‘basketball court.’ Copainalá Zoque *pecatza’ma* might be analyzable as /*peca*-*tzaman*/ ‘old-amate.tree.’ Guaraní *ka’aguy* appears to contain *ka’a* ‘mate herb,’ and Wayampi *kaʔa* also means ‘grass.’ Hawaiian *moku* can also refer to a ‘district,’ an ‘island,’ a ‘fragment of something,’ as well as, when used verbally, mean ‘to be cut, broken’ inter alia, while Rotuman *vao* also denotes a ‘net, fishing net.’ Sedang *chíu* also means ‘burned’ inter alia, *príu* also ‘hailstone, ice,’ *să* also “bow of kotro bow trap,” and Yay *dəŋ* also means ‘to pickle.’

## 27. The Gold

Representation: 63%

Motivated: 34.2%

Thereof Analyzable: 20.3%

Thereof Colexifying: 14%

Thereof by Contiguity: 17.7%

Thereof by Similarity: 7.9%

Recurrent associated meanings: yellow, metal, money, silver, red, gold ornament, diamond in cards, jewel, iron

Motivated terms for ‘gold’ have a clearly recognizable areal distribution. They are rare in the Old World and common in the New World, in particular North America, but are sometimes also found in South America and Eastern Eurasia, including Southeast Asia. A frequent pattern, found in ten languages (Japanese, Khalkha, Cheyenne, Chickasaw, Nez Perce, Oneida, Pawnee, Bororo, Maxakalí, and Rama), is to have complex terms for ‘gold’

where one of the constituents is a word for ‘metal,’ while in Nez Perce, Mandarin and Tetun, colexification of ‘gold’ with ‘metal’ is encountered. In analyzable items, the semantics of the second element is subject to some cross-linguistic variation. In Japanese, a term for ‘gold’ is *ō-gon* ‘yellow-metal’ (this structure is also found in Chickasaw, Kiliwa, Pawnee, where the relevant term colexifies ‘brass,’ Bororo, where the relevant term denotes any type of yellow metal generally, and Rama; in addition, note that Nez Perce has the optional complex term *maqsmáqs kícuy*, where *maqsmáqs* is ‘yellow’ and *kícuy* is ‘metal, money, dollar, gold.’ However, both constituents of the Nez Perce term can be used alone to refer to the concept ‘gold’). In Cheyenne, ‘gold’ is called *véhone-ma’kaata* ‘chief-metal;’ lexicographers note that this denomination may have come into being due to “the Indian chief head which has appeared on some old gold coins,” and note that this term may refer to gold coins specifically rather than the metal generally. Khalkha *erkim temyr* is analyzable as ‘supreme iron/metal,’ Maxakalí *pipkup yānām* as ‘metal.object light/lamp,’ Tetun *murak-mean* redundantly as ‘metal/gold-red/golden,’ Oneida *ohwistanolú* as /o-hwist-nolu-?/ ‘NOUN.PREFIX-metal-be.expensive-STAT,’ and Pawnee also has the term *paapicistaariksísu*?, which is analyzable as /paapicis-raariksísu-/ ‘metal/money-authentic-NOM.’ Both the Oneida and Pawnee terms in fact also mean ‘silver,’ and this is also the case in the Santo Domingo de Guzmán dialect of Pipil. In Lakhota and Malagasy, the relevant terms are analyzable (*mázaskázi* /mázaská-zí/ ‘silver-yellow’ and *volamèna* /vòla-mèna/ ‘silver-red’). In fact, associations with the color terms for either ‘yellow’ or ‘red’ are also commonly found and that in combination with elements of varied semantics, as already seen in some of the above examples; this is also common in Indo-European (Buck 1949: 610). Alongside the already explicitly mentioned cases of Japanese, Chickasaw, Pawnee, Bororo, and Rama with terms of the ‘yellow-metal’-type, as well as Lakhota and Nez Perce, an association with the color ‘yellow’ is found in other languages. Most often, the second element in terms of the analyzable lexical type is ‘money,’ as in Buli *ligmoaning* /ligra-moaning/ ‘yellow/red-money.’ Terms with constituents meaning ‘money’ and ‘yellow’ are also found in Kiliwa, Chayahuita, Miskito, Pawnee, and Rama (for the Nez Perce and Pawnee cases, compare discussion above). Terms with other or more specific patterns include Khoekhoe’s *lhuni-lurib* ‘yellow-iron,’ Tuscarora’s *tikačĩ?tkwáhnayē?*, which is analyzable as /ti-ka-čĩ?tkwahn-yēT/ ‘PARTITIVE-3SG.INDEF.AGENT-yellow-lay,’ and Guaraní’s *ita-ju* ‘rock-yellow.’ The association with the color ‘yellow’ is by colexification in Sahu, Quileute, Hani, and Vietnamese (it cannot be excluded that in some of these cases ‘gold’ is meant as a color term rather than the name of the metal). Alongside the languages already mentioned, the association with ‘red’ is also found in Upper Chehalis (*čšćíq t tá’la* ‘red INDEF.ART money,’ *tá’la* being a loan from English *dollar*), and Rotuman *monē mi’a* ‘money red.’ As these examples make clear, a frequent constituent for terms for the ‘gold,’ by functional contiguity, is ‘money.’ This is found in a total of eleven sampled languages, most of which have already been discussed: these are Buli, Biloxi, Upper Chehalis, Comanche, Nez Perce, the Santo Domingo de Guzmán dialect of Pipil (colexifying also ‘lead’), Pawnee, Miskito, Rama, Wayampi, and Rotuman (in addition, the Chayahuita term is semianalyzable and the recognizable constituent is a term for ‘money’). The association is by colexification in Comanche, Nez Perce, Pipil, Pawnee, and Wayampi. In line with this, Oneida *ohwistanolú*,



which also means ‘silver,’ is analyzable as /o-hwist-nolu-?/ ‘NOUN.PREFIX-metal-be.expensive-STAT,’ and Ancash Quechua colexifies ‘gold’ with ‘very valuable, precious.’

Finally, three sampled languages, Hausa, Badaga, and Bororo, by provenience or configurational contiguity, colexify ‘gold’ with ‘gold ornament’ or particular ornaments made of gold (Bororo also with a type of labret), and two further languages, Carib and Sedang, colexify ‘gold’ and ‘jewel.’ Wappo ?óro? and San Lucas Quiaviní Zapotec oor (both < Span. oro) also mean ‘diamond in cards’ (San Lucas Quiaviní Zapotec oor also means ‘time, hour’ due to phonological collapse of Span. oro and hora). Similarly, Basque colexifies ‘gold’ with ‘gold suit’ (in the Spanish Baraja deck of cards).

Further motivated and possibly motivated terms for ‘gold’ include: Efik *kut* inter alia also means ‘to see, to feel,’ Ngambay *tùngu* is also the name of a brown fruit, and Muna *bulawa* may be related to *bula* ‘albino, white skin.’ Sora colexifies ‘gold’ with ‘wealth.’ Comanche *puhihwi* is analyzable as /puhi-ekahwi/ ‘leaf-shiny,’ since this term also colexifies the meaning ‘money,’ a possibility is that this term originally referred to a banknote and was only later extended to ‘gold.’ Wintu *yo’qas* contains *yOq* ‘wash, pan for gold, gold dust,’ and Carib *ikamipo* is derived from *kami* ‘to glisten’ by the circumfix nominalizer *i-...-po*. Bislama *gol* also means ‘goal’ (< Engl. *gold* and *goal*), and Bwe Karen *thé* also ‘bear’ inter alia. Hawaiian colexifies ‘gold’ with ‘plain,’ ‘pasture’ and other meanings (the ‘gold’-reading of *kula*, the relevant term, is due to secondary English influence). Likewise due to the collapse of an English loan with native lexical material, *koro* also means ‘fortress’ inter alia. Finally, White Hmong *kub* also means ‘horn’ and ‘to burn,’ while Yay colexifies ‘gold’ with “loved or treasured person” and ‘needle.’

## 28. The Grass

Representation: 88%

Motivated: 37.2%

Thereof Analyzable: 2.9%                      Thereof Colexifying: 34.2%

Thereof by Contiguity: 19.0%              Thereof by Similarity: 0.8%

Recurrent associated meanings: hay/straw, meadow/pasture/lawn, plant, vegetable, green, leaf/foilage, (garden) rubbish, shrub, undergrowth, tobacco, fodder, forest

Lexico-semantic ties between ‘grass’ (or ‘weeds, herb’ – these are not distinguished for present purposes, and neither is ‘grass’ collectively as opposed to individualized ‘blade of grass’) and other concepts are typically realized by colexification, where the underlying semantic relation is one of contiguity. Globally, the most frequently found pattern (in 22 languages) is that languages feature a single term for ‘grass’ and ‘hay’ and/or ‘straw.’ This is particularly frequent in North America, but also figures in a significant number of languages in other parts of the world, namely in Buli, Ngambay, Abzakh Adyghe, Badaga, Bezhta, Khalkha, Kildin Saami, Upper Chehalis, Cheyenne, Chickasaw, Highland Chontal, Itzaj, Kashaya, Kiowa, Lakota, Nez Perce, Nuuchahnulth, Oneida, Pipil, Tuscarora, Yaqui, and Mandarin (as well as in Indo-European languages, Buck 1949: 519-520). Yahi and Central Yana *maadu* is also glossed as ‘wild hay.’ Turner and Brown (2004) interpret the presence of this pattern in Salishan and other languages of North America as a process of lexi-

cal acculturation: since ‘hay’ as a foodstuff for domestic animals was unknown in pre-contact times, there was no distinct term for it, and the native terms for ‘grass’ experienced semantic extension to also cover this aspect.

In fourteen further languages, Buli, Rendille, Kwoma, Ngaanyatjarra, Toaripi, Badaga, Greek, Khalkha, Welsh, San Mateo del Mar Huave (colexifying also ‘firewood’), Arabella, Chayahuita, Kaingang, and Tetun, terms for ‘grass’ are found that at the same time also may be used to refer to a conjunction of grass, that is, a ‘meadow,’ a ‘pasture,’ a ‘lawn’ or a grass-covered area in general (in Buli, this meaning is associated with the plural form of the term). This is a fine instance of configurational contiguity (also attested diachronically in Indo-European according to Buck 1949: 519).

Just a little less frequent, occurring in thirteen languages, namely Muna, Abzakh Adyghe, Badaga, Bezhta, Ket, Kildin Saami, Upper Chehalis (by a semianalyzable term containing an element meaning ‘grow’), Cheyenne, Ineseño Chumash, Embera (colexifying also ‘thicket,’ ‘vegetation,’ as well as ‘mount, mountain,’ the meanings are associated with different genders), Guaraní, Maxakalí, and Wayampi is colexification of ‘grass’ with ‘(green) plant’ (sometimes also ‘bush’ and other similar meanings), that is, colexification of two adjacent ranks in the ethnobiological taxonomy (the relevant Muna and Wayampi terms also mean ‘forest,’ while similarly Rotuman colexifies ‘grass’ with ‘plants and bushes growing wild’ *inter alia*). Embera and Fijian also colexify ‘shrub’ more specifically, and Kaluli and Rotuman also ‘undergrowth.’

In four sampled languages, Rendille, Ngaanyatjarra, Khalkha, and Tuscarora, the term for ‘grass’ is associated lexically with the color term for ‘green’ (it appears to be quite frequent for color terms for ‘green’ to be lexically related to terms for ‘grass,’ see e.g. Callaghan 1979 on some Miwokan languages, Proulx 1988 on Algonquian languages, and Buck 1949: 520 for diachronic evidence from Indo-European). In Tuscarora, there is also a complex term for ‘grass’ on the basis of the root *-her-* ‘green, grass,’ namely *uherúhkweh*, analyzable as /u-her-(a)hkw-eh/ ‘NOUN.PREFIX-green/grass-INSTRUMENTAL-NOUN.SUFFIX.’ This term has a wide semantic range, which includes alongside ‘grass’ also “ground, cover, hay, reed, rush” as well as ‘weed.’ Alongside these languages, there is a semianalyzable term for ‘green grass’ found in Welsh which appears to contain *glas* ‘green.’

Also in four languages, namely Efik, Swahili, Tasmanian (Middle-Eastern, South-eastern, and Western), and Xicotepec de Juárez Totonac, and here exclusively by colexification, an association with the meaning ‘leaf’ and/or ‘foliage’ is found (in Swahili also with ‘needle’). Other recurrent associations by colexification, some of which are particularly frequent in certain regions of the Old World include that with ‘(garden) rubbish’ (Baruya, Buin, Sora, Bora, in which latter also “cumbering things” is colexified), with ‘tobacco’ (Efik and Basque, which latter also colexifies ‘brow, forehead’), with ‘fodder’ (Badaga, Sora; an association very common in Indo-European, Buck 1949: 519–520), and with ‘vegetable’ (Efik, Ket, Khalkha, Kildin Saami, Guaraní, and Fijian, compare also Greek *chortári* ‘grass’ and *chortarikó* ‘vegetable’ and the colexification with ‘lettuce’ in Nez Perce).

Associations found only in one particular language only are many. Efik *i’kōñ* also means ‘bud,’ and Hausa *haki* also means ‘wages’ and may refer to “panting; gasping, breathlessness.” *Ciyawa*, another Hausa term, is also used to refer to the first signs of vari-

ous diseases. Khoekhoe colexifies ‘sheath of grass,’ and Noni *gayi* is also used to refer to so-called “elephant grasses” specifically. Rendille *roób* can also be used to characterize the fertility of vegetation, Angkor *wohi* also means ‘to weed,’ and Kaluli *kis* also may refer to ‘small weeds in garden.’ Nunggubuyu *maḍa* also refers to a fish species called ‘long tom’ which “swims with long, pointed mouth out of water, perhaps resembling a blade of grass.” Rotokas *rauritave kou* and *isiso kou* may contain *kou*, a classifier for ‘heaps’ according to Robinson (2011), and *isiso* is “steaming, giving off steam,” “steam something, heat with steam, scald.” Tasmanian terms in all varieties except the Western also denote ‘seaweed,’ while Yir Yoront *warrq* exhibits a typical Australian actual/potential polysemy, meaning both ‘grass’ and ‘grass or bush fire.’ Abzakh Adyghe *wac* is also can also mean ‘medicine’ (presumably originally made from herbs, i.e. by metonymy) and then, by further extension also ‘chemical medicine’ and ‘chemical product’ in general. Badaga *aṇe ~ haṇe* can also mean ‘grazing’ and ‘swamp, wetland,’ as well as “upper part of a ridge” and ‘water channel,’ Basque *belar* colexifies ‘forehead,’ and Khalkha ‘dish’ of food. Kildin Saami *rāss* also means flower,’ Upper Chehalis *smáqʷmumš* is analyzable as /s-máqʷm=umš/ ‘CONTINUATIVE-prairie=place,’ and Nuuchahnulth *ṣaqmapt* also denotes the ‘Brome grass’ specifically. For Kiowa *sųḍa* ‘plant or tuft of grass,’ compare *sųḍa* ‘grind; brush hair,’ and for San Lucas Quiaviní Zapotec *gyihzhyah*, compare *gyihah* ‘rock, stone.’ Tuscarora *unḗhsakwt katerʔahθétih* must be analyzable (*unḗhsakwt* may mean ‘by the house’), but the constituents cannot be identified with any certainty. Wintu *čaruq* ‘grass, greens’ also denotes ‘edible clover’ specifically. Arabela *jiya-socua* is one of the few cases of a complex term for the concept, literally meaning ‘land-CLASS.ROUND.TIMBER.’ Guaraní *ka’a* also denotes a “type of Paraguayan tea” and ‘vegetation’ generally. Bislama colexifies ‘grass’ with ‘feather,’ ‘pubic hair,’ ‘fern’ and ‘mould,’ Fijian with ‘to pierce’ inter alia, and Bwe Karen with ‘to command’ and other meanings. Hani *jahhaq* might be related to *haq*, meaning inter alia ‘bitter,’ Hawaiian *mau’u* is also the name for a ‘kava strainer’ and ‘strand of pandanus plaiting’ used in making hats, while Mandarin colexifies ‘grass’ with “disorderly, negligent.” Finally, Lesser Antillean Creole French *zeb* can also mean ‘black magic.’

## 29. The Headland

Representation: 26%

Motivated: 47%

Thereof Analyzable: 30.8%

Thereof Polysemous: 18.8%

Thereof by Contiguity: 10.3%

Thereof by Similarity: 29.5%

Recurrent associated meanings: nose, point/tip, head/forehead, end/to end, corner, sprout

‘Headland’ (or ‘cape, promontory’) is one of the meanings predominantly expressed by metaphor-driven analyzable terms. In fact, the most frequent association is a fine example of a body-part metaphor. This is the association with ‘nose,’ clear cases of which are found in six of the 39 languages for which data are available. This association is realized by analyzable terms in two languages, Basque and Vietnamese (*lur-mutur* ‘earth-snout/nose’ and *mũi đất* ‘nose land’ respectively), and by colexification in Kolyma Yukaghir, Haida (in

which the relevant term can also inter alia mean ‘end’ in general, for which compare colexification with ‘to end, be finished’ as well as “ultimate, final, last” and “hair, lady’s locks, tuft of hair, knot” in Badaga and Welsh *pen-rhyn* ‘end/head/top/mouth-hill’ and *pen-tir* ‘end/head/top/mouth-land’), Fijian (colexifying also ‘mountain peak’ inter alia), and Rotuman (which also has optional complex terms). In addition, the Ineseño Chumash and Great Andamanese terms are semianalyzable, and one of the identifiable constituents is the respective word for ‘nose.’ In four languages, another body-part metaphor is found, the source concept this time being the ‘head’ or ‘forehead,’ for which compare the etymology of Engl. *cape*, ultimately related to Latin *caput* ‘head’ (Oxford English Dictionary). The relevant languages are Yoruba (*iyorí ilẹ̀ ògògòrò*, analyzable as /i-ɔrí ilẹ̀ ògògòrò/ ‘NMLZ-to.sprout.or.appear head’), Welsh (*pen-rhyn* ‘end/head/top/mouth-hill’ and *pen-tir* ‘end/head/top/mouth-land’), Lavukaleve, and Hawaiian, where the association is by colexification (but in Lavukaleve it is more precisely ‘forehead’ which is colexified, and in Hawaiian also ‘brow’ and ‘wisdom’ inter alia are colexified). In addition, Swahili *rasi* assumes the meaning ‘head’ in poetic languages (this association is already present in Arabic, from which it is borrowed), and the case of Toaripi *harihari* is somewhat similar to this pattern, as this term also denotes the ‘temples’ (as well as ‘river bend’); the same goes for Efik, where *isü* in *isü iköt* can, alongside ‘tip of tongue, tip of finger’ and other meanings in a variety of contexts, also mean ‘face’ (*iköt* is ‘forest’). Two sampled languages, Tuscarora and Mandarin, have complex terms for the meaning ‘headland’ in which ‘corner’ acts as a source concept, for instance, the Mandarin term is *hai3-jiao3* ‘sea-corner/horn’ (the Tuscarora term colexifies ‘cave’ as well as ‘nook’). Another case of shape-based similarity with a body-part is found in Takia *damo-*, which also means ‘shin bone.’ Four languages’ terms betray associations not characterized by metaphor, in that the source concept is ‘point’ or ‘tip’: these are Efik (for data see above), Buin (*tiutiuna*, reduplicated from *tiuna* ‘point, peak of mountain, junction of rivers’), Ket (*solgup* /saʔl-küb/ ‘sharp-point’ for a “piece of land jutting out into river”), Kashaya (*?ama pʰišuša?*, containing *?ama* ‘land’ and *-?šuš-* ‘be pointed’), and Rotuman, where in fact ‘nose’ and ‘point, tip’ are both colexified with ‘headland’ alongside other meanings. Moreover, in Hawaiian one term for ‘headland,’ *oi’oina*, is a nominalization of *oi’oi* ‘pointed, protrude, sharp,’ and in Aymara, *moqo* may also refer to a ‘protuberance’ of any kind. The Tetun term for ‘headland’ is *rai-dilan* ‘land-bud/sprout/shoot,’ for the similar conceptualization in Yoruba see above.

Further patterns in the available data are: Efik *ukabare* is derived via the nominalizing prefix *u-* from *kabare* ‘to turn’ and also means “a turning around; varying, variation” as well as a ‘turning point’ and ‘turn’ generally. Buin colexifies ‘cape’ with ‘peninsula,’ and Rotokas *kipekipea* also means ‘point at base of ridge’ and ‘spine of lizard.’ Sko *pong* can also function as a verb meaning ‘to be closed’ and ‘to blow at fire,’ but this may be due to homonymy rather than a genuine case of lexical motivation. Chukchi colexifies ‘prow of boat.’ Nivkh *knyk* also means ‘rock, cliff,’ and Nuuchahnulth *?apquuʔa* is also a toponym denoting a specific point of land. Malagasy *tànjona* can also mean ‘aim’ or ‘objective’ and *tsiraka* also denotes a “sandy seashore,” while Samoan *tolotolo* is reduplicated from *tolo*, which has a large amount of different meanings, among them ‘to crawl,’ ‘to pull,’ ‘to mix,’

and also 'yam tuber,' which is perhaps the most likely reduplication base because of similarity in shape.

### 30. *The Honey*

Representation: 77%

Motivated: 53.2%

Thereof Analyzable: 27.7%      Thereof Polysemous: 25.4%

Thereof by Contiguity: 31.7%      Thereof by Similarity: 14.1%

Recurrent associated meanings: bee, liquid/sap/juice, sweets/sugar, beehive, syrup, fat/grease, resin, honeycomb, faeces, nectar, wax, jam, gum, oil, egg

'Honey' is an interesting concept cross-linguistically in that in some languages, the same term is actually used to denote both the product 'honey' as well as its producer, that is, the 'bee.' This pattern is found in 17 of the sampled languages, namely Hausa, (denoting "honey together with the comb"), Ngambay, Yoruba, Burarra, Ngaanyatjarra, Nunggubuyu, Yir Yoront (by a number of semianalyzable terms denoting different types of bee species and the honey they produce, all containing *may* 'vegetable food,' one of them also containing *ponh* 'hornet' and another one containing *kathn* 'yamstick'), Chickasaw, San Mateo del Mar Huave, Xicotepec de Juárez Totonac, Cavineña, Kaingang, Lengua, Miskito, Wichí, Yanomámi, and Hawaiian, and thus in all areas of the world except Eurasia, including Southeast Asia (compare the absence of the pattern in Indo-European, Buck 1949: 383-384); moreover, Bislama *honet* 'wasp, hornet' rarely assumes the meaning 'honey.' There are also languages with a variant of this pattern, namely languages in which the meanings 'beehive' and 'honey' are colexified. These are Gurindji, Itzaj, and Yanomámi, while in Bora and Chayahuita, the association with 'beehive' is realized by complex terms involving classifiers (*íimúhojpácyo* /*íimúho-hpácyo*/ 'beehive-CL.liquid' and *nino-i* 'beehive-CLASS.LIQUID' respectively; Bora *íimúho* 'beehive' is in fact itself analyzable as 'sweetness-CL.oblong,' and the language has the alternative term *iimúbá arááve-jpácyo* /*iimúbá arááve-hpácyo*/ 'sugar.cane condense-CL.liquid,' with *iimú-bá* 'sugar cane' analyzable as 'sweetness-CL.3d'). Moreover, Sora has *adaŋ'da:n* ~ *adaŋ'da:n*, presumably containing elements meaning 'beehive' and 'water, liquid,' and Cubeo has *mumicoro*, consisting of the root *mumi* associated with bees and the classifier *-coro* for liquid states, and in Yanomámi all three semantically related meanings – 'bee,' 'honey,' and 'hive' – are expressed by the same term *puu*. The same situation obtains in Nunggubuyu.

However, even more frequent than colexification with 'bee' are complex terms in which this meaning serves as a contiguity anchor. Within this class of terms, several recurrent patterns are encountered. By far the most common cross-linguistic tendency, encountered in the form of fully analyzable terms in ten sampled languages, Kyaka, Meyah, Rotokas (here, 'bee' and 'honeycomb' are colexified, while in Nunggubuyu and Kiliwa, 'honey' and 'honeycomb' are, in Kiliwa by the analyzable term *mi?-yaw-y* 'light=larvae-ATT'), Toaripi, Pawnee, Yaqui, Maxakalí, Piro, Rama, and Tetun, is to have terms for 'honey' consisting of the word for 'bee' (or rarely 'wasp') and 'water,' 'liquid,' 'sap,' or 'juice.' Note also that Hausa, which is of the colexifying type, also features the

compound *ruwa-n zuma* ‘water/liquid-GEN honey/bee,’ denoting “honey apart from the comb” specifically; the situation in Miskito and Hawaiian is comparable. However, notably, even this association is not exclusively realized by morphologically complex terms. In Nunggubuyu, *-(w1)arḡayag* (which may be related diachronically to *-ar-* ‘water’) also means ‘juice,’ Basque *ezti* means both ‘sap, juice’ and ‘honey’ (alongside ‘graft, grafting’ and ‘vaccine’ inter alia), and in Huambisa, *yumiri* is used with reference to both ‘liquid’ in general as well as ‘soup,’ ‘juice,’ and ‘honey’ specifically. There are redundant terms with this structure on the basis of the term colexifying ‘bee’ and ‘honey’ in Yanomámi and Hawaiian. A semianalyzable term where the identifiable constituent is ‘water’ is found in Berik.

In four languages (Efik, Mbum, Biloxi and Carrier), an association by complex terms with ‘fat’ or ‘grease’ is found; for instance, in Carrier, ‘honey’ is *t̥sihna-rě* ‘wild.bee-grease.’ In addition, in Nunggubuyu, the association is realized by colexification, and Ngambay has a complex term of this kind which is however redundant due to colexification of ‘bee’ and ‘honey.’ In two languages, Efik and Kwoma, an association with ‘oil’ is found, for instance in Kwoma *heemi kwar* ‘bee oil’ (*kwar* is more precisely “the name of two related types of tree with very oily wood;” the fact that Efik figures in both categories is because the constituent element in fact can refer to either ‘fat’ or ‘oil’). A semantic pattern either realized by analyzable terms or by colexification is the association with the meanings ‘sweet(s)’ or ‘sugar.’ Next to the somewhat different case of Bora already discussed above, it occurs by means of morphologically complex terms with the other constituent being ‘bee’ in Upper Chehalis, Comanche, and Hani (for instance, in the latter language one of the terms for ‘honey’ is *biaq-qul* ‘bee-sweet;’ in Comanche, *ʔnʔ bihnaa* is analyzable as *ʔnʔ pihnáaʔ* ‘insect sugar’ more specifically), and by colexification in Muna, Basque, San Mateo del Mar Huave, Kiowa (colexifying also ‘panocha’), Lake Miwok, Yaqui, Aymara, Guaraní, Hupda, and Imbabura Quechua. Xicotepec de Juárez Totonac colexifies ‘honey’ with ‘sugar cane conserve,’ Pipil (Cuisnahuat dialect) has *-nekayut*, containing *nek* ‘candy, sweets’ and *ayut* ‘liquid, juice.’ In addition, the semianalyzable Abzakh Adyghe term *śew* contains *śe* ‘sweet, sugar,’ and redundant terms with this structure on the basis of a simplex colexifying ‘bee’ and ‘honey’ directly are encountered in Chickasaw and Miskito. A similar pattern predominantly found in North America (Upper Chehalis, Cheyenne, Itzaj, Nez Perce, Oneida), but also in White Hmong in Southeast Asia is that with ‘syrup.’ For instance, in Upper Chehalis, *cíčsməlās* is analyzable as */cíčs-məlās/* ‘bee-syrup,’ *məlās* being a loanword from French (< *mélasse*) that entered the language via Chinook Jargon. The association is by colexification in Itzaj, Nez Perce and Oneida.

Haida *sraal narii* is analyzable as */sraal nara/* ‘bee faeces.’ Other languages with terms with such structure are Blackfoot and Fijian. In one sampled language, Buin, the same term is used for ‘urine’ and ‘honey’ (the relevant term is also a toponym for a river and a mountain), a pattern which one might have expected to be more frequent given that, alongside the tertium comparationis available for both ‘faeces’ and ‘urine,’ namely that all three are excretions of a living being, here also the fact that both ‘honey’ and ‘urine’ are fluid is available. In Dadibi and Bislama, an association by morphologically complex terms with ‘egg’ is found; in Dadibi ‘egg’ is colexified with ‘nut’ and ‘round object generally,’ and in Bislama, the term denotes especially honey “taken directly from the

hive,” while otherwise *hani* (< Engl. *honey*) is also used. Wappo is unique in the sample in that in this language, ‘honey’ is called *tuš-huméy* ‘bee-tear.’ There is a semianalyzable term where the identifiable constituent denotes a certain bee species in Piro, and one with the generic term for ‘bee’ in Hani.

By colexification, three sampled languages associate the meanings ‘honey’ and ‘resin.’ These are Nez Perce, Oneida, and Bwe Karen (colexifying ‘to be pure, clean,’ *inter alia*), and Tuscarora has *ruʔtáhkə uθrəwəh*, consisting of *ruʔtáhkə* ‘bee’ and *uθrəwəh*, which can refer to ‘resin,’ but also ‘cement,’ ‘glue,’ ‘gum,’ ‘jam, jelly,’ ‘molasses,’ ‘syrup,’ ‘tar,’ and ‘wax’ (‘jam’ and ‘honey’ are colexified directly in Gurindji, and ‘gum’ and ‘honey’ are in Oneida; see above for the association with ‘syrup’). As is clear from this list, this pattern appears to be particularly frequent in the Americas. Buli colexifies ‘honey’ and ‘wax’ (and this may be the case in Nunggubuyu, too, but this is not sure), and Khalkha and Vietnamese share colexification with ‘nectar.’

Further unique associations in the sample include the following: Koyraboro Seeni *ayuu*, a dialectal variant of *yuu*, also means ‘manatee,’ and for Ineseño Chumash *ʔaqʔikaš* compare *aqʔikaš* “to burn in the throat from sweetness, as honey does.” Nunggubuyu *mala* also means ‘navel’ and ‘(clear) sky,’ Ket colexifies ‘honey’ with ‘copper,’ and Nez Perce *temísquy*, alongside ‘syrup’ and ‘honey,’ also means ‘sorghum’ and denotes the “gelatinous sap of the tamarack tree.” Bora, alongside a term featuring an association with ‘beehive,’ also features other terms formed with the classifier for liquids, *-hpácyo*, namely *ócóoméhojpácyo*, where *ócóomého-* denotes the hive of a particular type of bee. The Wayampi term *ei ~ ey* also can refer to a religious ‘honey dance’ that the Wayampi perform. Manange *4k<sup>we</sup>* also means ‘song,’ and Hawaiian *hone* (at least in the sense of ‘honey’ < Engl. *honey*) colexifies ‘honey’ with “sweet and soft” (said of music) and “sweetly appealing” (said of perfume) and other meanings. Furthermore, there are a number of terms in different languages in which the term for honey strikingly resembles semantically related words, but which resist morphological analysis. These include the following pairs: Highland Chontal *algujua* ‘honey’ and *galgujua* ‘honey bee,’ Wintu *huʔbi* ‘honey of bumblebee’ and *huʔbit* ‘yellowjacket’ (sharing the root *huʔb-*), Huambisa *yumiri* ‘liquid, soup, juice, honey’ and *yumin* ‘sweet,’ and finally Lengua *yohena* ‘honey’ and *yohan* ‘bee.’ Finally, it should be noted that several sampled languages have a more specific honey terminology, in which one general term stands alongside a range of more specific terms for different types of honey, and in Arabela, a general superordinate term is lacking altogether and only specific types of honey are lexically designated.

### 31. The Horizon

Representation: 28%

Motivated: 66.5%

Thereof Analyzable: 67%

Thereof Colexifying: 1.9%

Thereof by Contiguity: 3.7%

Thereof by Similarity: 54.9%

Recurrent associated meanings: sky, edge/border/fringe, end/finish, land/earth, meet/meeting place, cloud, basis, sea

As the above figures make clear, the ‘horizon’ is one of the concepts predominantly expressed by analyzable terms driven by metaphor. The most frequent lexico-semantic association is that with meanings such as ‘edge,’ ‘border,’ or ‘fringe.’ This is found by colexification in Buin (though the relevant term *kitai(na)* appears to be derived from *kita* ‘split longitudinally, split with the grain’) and by analyzable terms with ‘sky’ acting as contiguity anchor in Abzakh Adyghe (there is also an alternative version of the term where ‘sky’ is replaced by ‘cloud’), Basque, Khalkha, Carrier, Haida (where the term can be literally translated as “(where) the edge of the sky hits the land or water”), Bororo (which also has the alternative term *barubaru*, reduplicated from *baru* ‘sky’ and colexifying ‘type of spirit’), Kaingang, and Hani (where *aoqkaq milkaq* contains *aoq* ‘sky’ and *milkaq* ‘border line,’ there also is another term with similar structure featuring *zeil* ‘line’). There is a derived term in Hawaiian which also denotes the “[c]ords or fine ropes threaded through marginal meshes of upper and lower edges of nets” and also means ‘trickery, deceit, deceive, cheat,’ and similarly, Greek *orízōn* is related to *orízō* ‘to bound, delimit.’

A very similar conceptualization strategy where ‘end’ is the source concept is found with ‘sky’ as the contiguity anchor in Efik and Guaraní (*utit enyöñ* ‘end sky’ and *arapaha /ára-paha/* ‘sky-end’ respectively). Similarly, Maxakalí has *hāmnōgnōy*, presumably containing *hām* ‘earth’ and *nōg* ‘finish off, use up.’ In Buin, the association is by colexification; it is the only sampled language to behave in this way. In three languages, Upper Chehalis, Miskito, and Hawaiian, an association with verbs meaning ‘to meet’ or nouns meaning ‘meeting-place’ is found. For instance, Upper Chehalis *xʷuqʷxʷuqʷəlwtstn* contains *xʷúqʷi-* ‘to gather, come to be gathered’ and the lexical affix *=tn* ‘place.’ In Hawaiian, this association is by colexification. Furthermore, in three sampled languages of Oceania, Kosarek Yale, Fijian, and Hawaiian, analyzable terms featuring an element meaning ‘basis’ are found (Fijian *vū-ni-lagi* ‘basis-poss-sky,’ Hawaiian *kumu-lani* ‘base-sky,’ Kosarek Yale *imbubuak* and *soobubuak*, containing *im* ‘sky’ and *soo* ‘earth, soil’ respectively alongside *bubu* ‘liver, carrying part, point of support’). Also in three sampled languages, one of the constituents of the term for ‘horizon’ is that for ‘cloud.’ One of the Abzakh Adyghe terms for the ‘horizon’ is, as alluded to above, *pšape* /pše-pe/ ‘cloud-edge/tip,’ in Biloxi, it is called *nací’-xwūhi* ‘cloud-low’ (this term indeed also denotes a ‘low cloud’), and in Itzaj *chäkil muyal* ‘red cloud.’ As has already become clear from the previously mentioned Kosarek Yale and Haida terms, in some languages a constituent of the word for the ‘horizon’ is a term with the meaning of ‘land’ or ‘earth.’ This is also found in Maxakalí and Miskito. Further terms where one constituent is ‘sky’ are Ngambay *gél-dàra* ‘origin-sky,’ Khalkha *tngri jin xormai* ‘sky GEN lower.hem/foot.of.mountain,’ Kiliwa *?+ma?i=haa=p+?iw-m-u?* ‘DN+sky=move/go=PASS+stand-THR-OBL,’ a more palpable literal translation of which is “where the sky stops,” Tuscarora *weyurehya?nihe* (built around the roots *-rehy-* ‘sky’ and *-?niha-* ‘sprain’), Bororo *barubaru* (reduplication base: *baru* ‘sky,’ the complex term is also the name for a kind of spirit), Hani has *aoq-daoq aoq-zeil*, containing *aoq* ‘sky, heaven’ and *zeil* ‘line,’ Hawaiian *pō‘ai-lani* ‘circle-sky,’ Lenakel *noua-nisii-neai/tehe* ‘fruit-excrement-sky/sea’ (for which compare Carib *palana lali* ‘sea floor’ and Hawaiian *‘ili-kai* ‘surface-sea,’ which indeed also denotes the ‘surface of the sea’ and ‘horizontal’ more abstractly), and



Rotuman *lā ne laqi* ‘foot ART.PL sky.’ There are semianalyzable terms in Central Yup’ik, Cubeo, and Guaraní.

Further associations include: Khoekhoe *mûs (di) lkhâu-s* is analyzable as ‘eye (POSS) radiate-3SG.FEM’ (Khoekhoe has also borrowed *horisonni* from English or German), Nuuchahnulth has a lexical suffix colexifying ‘on the horizon’ with ‘far out at sea,’ and Chayahuita *hui'tontarinso* contains *hui'tonin-* ‘paint with horizontal strokes’ (for which compare colexification of ‘horizontal’ as well as ‘surface of the sea’ in Hawaiian), and there is another term, *aquë notëhuatëra*, containing *aquë* ‘far.’ Maxakalí *hăpkux-yă* may be analyzable as ‘shoreline-fragment,’ for Hani *caqqiq*, compare *caq* ‘a mark’ and *qiq* ‘to lift up’ (both possible constituents also have further meanings). Hawaiian, featuring a wealth of terms for the concept, also has *Kahiki moe* ‘Tahiti prostrate,’ as well as *‘alihi lani*, which is redundantly analyzable as ‘horizon sky’ and colexifies “deceit, trickery” *inter alia*. Samoan *i tafatafa-‘ila-gi* is analyzable as ‘PART side-spot-SUFFIX’ and means ‘on the horizon’<sup>4</sup>

### 32. *The Horn*

Representation: 86%

Motivated: 44.7%

Thereof Analyzable: 7.2%

Thereof Colexifying: 37.5%

Thereof by Contiguity: 12.2%

Thereof by Similarity: 24.3%

Recurrent associated meanings: antler, musical instrument, branch, tusk, antenna, point, ear, tooth, cockscomb, bone, post, tributary, clavus, nail, thorn, signal horn, corner, wood

Analyzable terms for the meaning ‘horn’ are relatively rare cross-linguistically, and where they occur, it is typically in areas without indigenous species of large horned animals, such as Australia and the Amazon region of South America. An association with ‘ear’ is found by complex terms in Jarawara (*warabi awe/warabi ewene /narabi/narabo awe/ewene/* ‘ear piece.of.wood,’ for which compare Kaingang *nĩ-ka* ‘flesh-tree/wood’), Wayampi (*ămilāsĩ*, containing *nami* ‘ear’ and *āsĩ* ‘pointed’), and Hawaiian (*pepeiao-hao* ‘ear-iron,’ with *hao* also being the name of the horn of a goat itself). Similarly, Yir Yoront has *pin+ngon* ‘ear+hornlike.process;’ in addition, the association is realized by colexification in Burarra (where the relevant term may also refer to any “appendage that sticks out like an ear” as well as “witness to what was said”). Great Andamanese has *wôlo-tâ* ‘adze-bone.’ This term denotes the ‘horn of cattle’ specifically, and the translation of the example for its usage provided in the source makes particularly clear that it is a neologism for a newly encountered concept: “when we first saw cattle we called the horns (lit. things on their heads) *wôlo-tâ (da)*, i.e. adze(-like) bones.” By colexification, the association is found in Mesoamerica, namely in Highland Chontal and Itzaj (in the latter language, *b’ak*, the term in question, also means ‘employment;’ the association with ‘bone’ is also noted for Kolyma

<sup>4</sup> The source for Yoruba also lists *ibiti o dabi enipe ilẹ ati orun pade* ‘place:REL 3SG appear like land and sky meet;’ it is not mentioned in the above since Joseph Atoyebi (p.c.) informs that this is not a conventionalized part of the Yoruba lexicon.

Yukaghir according to an older source incorporated into the consulted source). In fact, the association with ‘ear’ and ‘bone’ is the only recurrent one typically realized by morphological analyzability. Further analyzable terms include Central Yup’ik *ciru-neq* ‘cover-area,’ Bora *iyáábej-to* ‘animal-CL.spine,’ and Kapingamarangi *madaagoo* /mada-goo/ ‘end-husking.stick/coccyx.’

Associations by colexification, in contrast, are much more varied cross-linguistically, although also here recurrent tendencies emerge. Basque, Chukchi, Nivkh, Kolyma Yukaghir, Carrier, Upper Chehalis, Cheyenne, Haida, Kiowa, Lake Miwok, Nez Perce, Quileute, Wappo, Wintu, Yuki, Central Yup’ik, Tehuelche, Bwe Karen, Hani, Hawaiian, and Sedang colexify ‘horn’ with ‘antlers.’ In five languages, Rotokas, Chickasaw, Ancash Quechua, Wayampi, and Tetun, the word for ‘horn’ is also used with the meaning ‘antenna’ or ‘feeler’ of an insect, or, according to the Ancash Quechua source, of a ‘worm’ in this language (compare also Sora *dereṇən* ‘horn’ and *dereṇbudən* ‘antenna of insect’), while six languages, Hausa, Ngambay, Basque, Kildin Saami, Abipón, and Toba, have one term covering both ‘horn’ and ‘branch’ or similar meanings (Basque also colexifies ‘tributary’ – a pattern also found in Toba- ‘lineage,’ and ‘bedpost, chairpost’). Furthermore, in three languages, Arabela, Ancash Quechua and Bislama, the same term is used for ‘horn’ and ‘cockscorn’ (in Bislama apparently also for ‘comb’), and in two, Khalkha and Lesser Antillean Creole French, ‘horn’ has a semantic extension to ‘clavus.’ In two further sample languages, Wintu and Fijian, an association with ‘nail’ is present (in Fijian by the analyzable term *i vako* ‘DERIV put.nail.through,’ while in Wintu, the same term may also be used to refer to a ‘hammer’ and a ‘drill’). Also exclusively by colexification, in three languages, Kyaka, Chukchi, and Samoan, ‘tooth’ and ‘horn’ are lexically associated (the Kyaka term also colexifies ‘food, edible material’ as well as “sharp, biting, erosive,” and the Samoan term *nifo* is also used figuratively to refer to an ‘enemy’). Similarly, Takia *fai* colexifies ‘horn’ with “upper canine teeth” specifically, and also with ‘crocodile.’ Kyaka (where the relevant term also has still other meanings) and Guaraní colexify ‘horn’ with ‘thorn.’ Similarly, Lake Miwok *kílli* ‘horn, antler’ seems to be related to *kíli* “to hook with the horns, thorn, stickers on weeds or berry bushes.” In addition, Bora and Guaraní colexify ‘horn’ and ‘point’ (compare Wayampi’s term *amilāsī* containing *āsī* ‘pointed,’ and the similar connection with ‘point, peak’ in Irish mentioned by Buck 1949: 208), while Aymara and Basque colexify ‘horn’ and ‘post’ (‘bedpost,’ ‘chairpost’ more specifically in Basque). For six languages, Kanuri, Basque, Welsh, Chickasaw, Lesser Antillean Creole French, and Hawaiian, lexicographers note that the word for horn is also used for a musical instrument, as is the case in many European languages (Welsh *corn* also colexifies ‘corn,’ ‘stethoscope,’ and ‘might’). Furthermore, in Khoekhoe, the same roots associated with different nominal designants yield the relevant meanings, and Quileute *pòxʷó-kʷot* may be conceived of as variation of the association with musical instruments reported above, as it also means ‘horn of a ship.’ Similarly, the relevant Aguaruna and Huambisa terms (both apparently borrowed from Spanish) may also refer to a ‘signal horn,’ Bislama colexifies ‘loudspeaker,’ and Lesser Antillean Creole French colexifies ‘motor horn,’ alongside ‘corn’ and ‘horny area.’ Finally, in six languages, Noni, Swahili, Baruya, Toaripi, Lenakel and Samoan, the word for ‘horn’ may also refer to a ‘tusk’ (as of a pig or an elephant, for instance), and Hani

and Mandarin colexify 'horn' with 'corner,' among other meanings (see Matisoff 2004: 352 for further evidence from both within and outside Southeast Asia).

As already mentioned, there are many different associations by colexification, and a lot of them are found in only one sampled language. Hausa *k'aho* (presumably by provenience contiguity) also denotes a 'cigarette-holder,' but also "a kind of tattooing" and "a person who is either destitute or gives nothing away." Ngambay *gàjì*, alongside 'branch' and 'horn,' also means 'rubbish' inter alia. Baruya colexifies 'horn' with "curved moon, crescent moon" as well as "curve below the navel where skirts fit." Gurindji is the only language in the sample that uses the same term for 'horn' and 'root,' whereas Muna uses the term *tandu* not only to denote 'horn,' but also verbally with the meaning 'to butt, gore' inter alia. Abzakh Adyghe *bž(e)* can inter alia also refer to a drinking vessel such as a glass or cup, probably originally motivated by the former function of horns as drinking vessels, Badaga *kombu* also may refer to a "horn shape made from hill pavement which is put inside the rooves [sic!] of houses after the uppaṭṭuva," and Greek *kéras* is also used to refer to a 'wing of an army.' Biloxi *ahi' ~ ahe' ~ ahé' ~ he* is unusual in having the semantic range of "skin, fingernails, horn, hooves, scales of fish, bark of tree," and Upper Chehalis *wináw*, by perceptual similarity, also denotes a 'wedge.' Kiowa *guqdei* also means 'afterbirth.' The Santiago Mexquititlan Otomí term *ndäni* also denotes a 'ram,' whereas Pawnee's *paariiku'* clearly showcases meaning extension by metaphor: it also means 'colon' and 'banana.' Arabela *tuhuaja* is also used to refer to 'adornment' as well as 'crown.' *Kácho*, the Embera term for 'horn' can also mean 'slice' or 'piece' (it is in fact borrowed from Latin American Spanish *cacho* with the same semantic potential). Guaraní *tatí* also means 'prow,' while Ancash Quechua *wagra* can also refer to a 'useless thing.' Yanomámi *yōra* not only means 'horn,' but also 'fin.' The motivational history for Fijian *i leu* appears to be complex: *leu* is "to extract, as a thorn from flesh" and the derived nominal *i leu* in the first place means "the instrument for so doing: tweezers," and then also 'a small stick or needle' and finally, presumably by yet another metaphorical extension, also 'the horns of an animal.' A similar line of meaning extension may be hypothesized for *seru-na*, which is analyzable as 'to.comb-POSS.' Hawaiian *kiwi* also denotes the "horn" of the Kala fish (known in English as the Bluespine Unicornfish, *Naso Unicornis*), and furthermore any curved or bent object, and indeed, in an adjectival sense, also 'curved' and 'bent.' Rotuman *ipesi* also denotes a particular type of 'flat wedge-shaped wooden spatula,' Samoan *seu* also means 'to steer,' 'stir, mix,' and 'to intercept,' White Hmong *kub* also means 'gold' (in this sense borrowed from Chinese) and 'to burn,' and, finally, Yay *kaw'* also is used to refer to "a vine or creeper" and a 'protrusion' generally.

### 33. *The Lagoon*

Representation: 38%

Motivated: 51.5%

Thereof Analyzable: 23.5%

Thereof Colexifying: 28.9%

Thereof by Contiguity: 16.7%

Thereof by Similarity: 30.4%

Recurrent associated meanings: lake/pond, water/liquid, sea, puddle, swamp, tide, reef, round, big

Commonly, languages do not differentiate lexically between ‘lake, pool’ and ‘lagoon,’ using a single term for both referents. This is the case in Kwoma (colexifying also ‘waterway, canal’), Tasmanian (Middle-Eastern and Southeastern), Toaripi, Yaqui (by a semianalyzable term containing an element meaning ‘water’), Copainalá Zoque, Arabela, Aymara, Bora, Chayahuita, Cubeo (by the analyzable term *maca-jitabũ* ‘faeces-puddle;’ compare colexification of ‘lagoon’ and ‘puddle’ in Copainalá Zoque and Ancash Quechua), Embera (where the meanings are associated with different genders), Guaraní, Huambisa, Miskito, Ancash and Imbabura Quechua, Tsafiki, Fijian, and Takia, while Greek has *limnothálassa* /límn-o-thálassa/ ‘lake-STEM.FORMATIVE-sea,’ Carrier ‘*a-pen-ket*’ ‘fog-lake-on,’ Hawaiian *loko kai* ‘lake/inside sea,’ and Mandarin *xie4-hu2* ‘pour.down/fall-lake’ and *jiao1-hu2* ‘reef-lake.’ Hawaiian has, like Mandarin, an association with ‘reef’ (*kai kohola* ‘sea reef.flats,’ which indeed can also refer to the ‘shallow sea within a reef’). Greek, Mandarin, and Hawaiian are not the only languages associating ‘lagoon’ with ‘sea’: Nivkh has *kerq ṇalu* ‘sea bay,’ there is a semianalyzable term in Welsh, and San Mateo del Mar Huave and Kapingamarangi colexify the meanings directly (along with ‘tide, salt water, salt’ and ‘to close up, shut’ in the former language; for the association with ‘tide,’ compare Samoan *tai-tafola* ‘tide-spread.out’ and that Rotuman *maka* is glossed as “tidal flat, wide stretch of beach covered only at high tide” and also means ‘to sing, dance’). Likewise, Yaqui is not alone with its association with ‘water’ or ‘liquid’: Yanomámi has *mono u* ‘round.bodied liquid,’ Takia parallelly *you i-lanti* ‘water 3SG-be.round,’ Bororo *kuruga*, perhaps analyzable as *kuru-ga* ‘liquid-go,’ and Tsafiki *hua pipilú*, containing *hua* ‘big’ and *pi* ‘water.’ Other complex terms involving a constituent meaning ‘water’ are San Mateo del Mar Huave *waj-yow* ‘neck/mane-water,’ denoting a specific lagoon, and Tetun *bee-lihun* ‘water-dam/tank.’ Moreover, Guaraní directly colexifies ‘water’ (and ‘river’) and ‘lagoon,’ there is a semianalyzable terms with a constituent meaning ‘water’ in Guaraní, and a further semianalyzable term with a constituent meaning ‘accumulation.’ Kwoma and Yanomámi colexify ‘lagoon’ with ‘swamp’ (Yanomámi *wawëwawë* is a reduplicated form of *wawë* ‘wide, empty’); Tetun *kolan* is also glossed as ‘saltwater swamp’ in parentheses. Bora colexifies ‘lagoon’ with ‘parched arm of river,’ and, conversely, Yir Yoront with “stretch of river where water remains in the dry season” as well as ‘billabong.’

Other associations include: in Buin, the relevant term also colexifies the meaning “rainwater hole containing brackish water,” Yoruba *òsà* also has a temporal meaning ‘space of time, season, interval,’<sup>5</sup> and Kwoma *naba* can also refer to a swamp or a canal. Japanese *kata* also means ‘shoulder’ but has different prosodic structure from *kata* ‘lagoon.’ Bororo features two terms with unclear motivations for the concept: *baru-bo* is analyzable as ‘sky-division’ and *kuru-ga*, as mentioned above, perhaps as ‘liquid-go.’ Finally, Fijian *totobu*, meaning also ‘deep place in center of stream,’ is presumably reduplicated from *tobu*, meaning “pool in a river, bathing hole, well” *inter alia*, and Hawaiian *kua-‘au* ‘basin inside the reef, lagoon’ is tentatively analyzable as ‘back/windward-project.’

<sup>5</sup> The consulted source also has *adágún ibití omi òkun nṣàn sí* which is not a conventionalized Yoruba lexical item according to Joseph Atoyebi (p.c.).

34. *The Lake*

Representation: 82%

Motivated: 45%

Thereof Analyzable: 14.6%

Thereof Colexifying: 31.7%

Thereof by Contiguity: 17.4%

Thereof by Similarity: 10.6%

Recurrent associated meanings: puddle, lagoon, water, river, swamp, sea,  
spring/well, dam, big, round, rain, bay, inland, stand/stagnant

As discussed in § 6.2.2.5., bodies of water are sometimes not lexically distinguished from the substance ‘water,’ and this situation is also found for ‘lake’ (or ‘pond’) in five of the sampled languages, Berik, Itzaj, Bororo (by a register-specific term), Jarawara and Bislama (and is attested diachronically in Indo-European, evidenced by cognates meaning ‘lake’ in Avestan but ‘water’ in Sanskrit, Buck 1949: 38). In the case of Buli, Hausa, Ngambay, Rendille, Berik, Badaga, Khalkha, Bororo (by a term largely restricted to ritual language), Guaraní, Maxakalí, Jarawara, Piro, and Bislama ‘river’ is colexified (‘perennial river’ more specifically in Rendille), while there are, due to colexification of ‘water’ and ‘river,’ terms betraying this association by analyzable terms in Efik, Sko, Tsafiki, and Kaingang. Moreover, there is a semianalyzable term where the identifiable constituent can refer to both ‘water’ and ‘river’ in Kosarek Yale. In Itzaj and Jarawara, ‘rain’ is colexified (and similarly, ‘rain water’ is in Comanche). The Jarawara term *faha*, like Bislama *wota* and Itzaj *ja*, is very general, the reference of which may include ‘water’ as a substance, ‘river,’ as well as ‘lake’ (and ‘juice’ and other meanings in Bislama).

More frequent are, however, associations with ‘water’ by way of morphologically complex terms. One pattern found in three languages of the Americas, Blackfoot, San Mateo del Mar Huave, and Tsafiki, is that the second element is ‘big,’ e.g. Blackfoot *ómahksíkimi* /omahksi-ikimi/ ‘older/large-liquid’ (note, however, that the San Mateo del Mar Huave term *nadam yow ngo mawaag* contains the unknown element *mawaag* plus, apparently, the negating morpheme *ngo*, and compare also the redundant complex term *tòku ba* ‘big water/lake’ in Ngambay, and the same situation in Itzaj). In three sampled languages, Yanomámi, Rotuman, and Takia, the additional element is ‘round,’ e.g. Yanomámi *mono u* ‘round.bodied liquid’ (compare in addition Lengua *yakyengyiam* ‘lake’ and *yakyeyi* ‘be round’). Other complex terms for ‘lake’ where one of the constituents is ‘water’ include Efik *mkpö-diök'hö-möñ* ‘thing-be.placed water’ and (*ëbiët*) *udiökhömöñ* /ëbiët u-diök'hö-möñ/ ‘(place) NMLZ-stand.level water,’ Sko *pato* /pa-tó/ ‘water-inside,’ Abzakh Adyghe *psə-wəc°ə-ye* ‘water-stand-PAST,’ Cheyenne *tsé-sééha mähpe* ‘that.which.is-spread.out water,’ Tetun *bee-lihun* ‘water-dam/tank’ (‘lake’ and ‘dam’ are colexified in Badaga and Chickasaw) and Rotuman *tən häe* /tənu häe/ ‘water contain’ (this term also means ‘puddle’ and ‘bay, inlet,’ which latter association is shared with Nez Perce). Cahuilla *pāl múyeqalet* contains *pāl* ‘water, river’ and *-múye-* ‘fill up;’ the literal translation provided by the lexicographer is “water which fills up.” In addition, Maxakalí *puxhep* contains *pux* ‘pour out’ and *hep* ‘blood, sap, liquid,’ and *kōnāgkox* is analyzable as /kōnā’āg-kox/ ‘water-hole.’ Furthermore, there is a wealth of semianalyzable terms for ‘lake’ where the element ‘water’ is discerni-

ble, but the full morphological structure cannot be elucidated. This is the case in Noni, Kosarek Yale, Kiowa, Comanche, Yaqui, Yuki, and Samoan.

Associations with the 'sea' are also found, either by colexification (seven languages, Ngambay, Khalkha, Nez Perce, Tuscarora, Tehuelche, Yanomámi, and Hawaiian; diachronic shift is well-attested in Germanic, Buck 1949: 38) or by morphologically complex terms. This is found in the Japanese term for a 'big lake' (*mizu-umi* 'water-sea'), but is also found in two languages of Africa which have interestingly a very similar conceptualization strategy. In Bakueri, the 'lake' is called *mmána mmánja* 'child sea' or *ngúma mmánja* 'younger sea,' and in Khoekhoe *hurirob*, a term used in Bible translations, consists of *huri* 'sea' followed by the diminutive suffix *-ro* and the nominal designant *-b*. Moreover, there is a semianalyzable term in Hani where the identifiable constituent means 'sea, ocean' as well as 'to soak in water.' In connection with the metaphorical transfer of a 'child' as a prototypically small referent (Jurafsky 1996), the Waris term *polomb* deserves some discussion. It obviously contains *pol* 'liquid,' and the second element looks like a truncated form of *ombol* 'son.' There is a parallel for this from the same area: Baruya features the term *budaqya* for a 'small lake' which is formally redundant, since the constituent *bunya* alone can already be used to refer to a lake, but the other constituent of this compound is *taqya* 'girl.' A further analyzable term is Yanomámi *wawëwawë*, the reduplication base of which appears to be *wawë* 'wide, empty'

Exclusively by colexification, associations between 'lake' or 'pond' and 'swamp' are found in eleven of the sampled languages, namely Buli (where the interpretation of the relevant term as 'lake, pond' is rare), Ngambay, Gurindji (where the term also denotes a specific lake and "something in ceremony"), Kwoma, Basque, Ket, Chickasaw, Lesser Antillean Creole French, Kaingang (by the analyzable term *óré ki goj* 'mud in water/river'), Pawnee, Yanomámi (by the term *wawëwawë*, reduplicated from *wawë* 'wide, empty'), and Hani (and see Buck 1949: 38 for Indo-European evidence), with 'lagoon' in 21 languages, Kwoma, Tasmanian (Middle-Eastern and Southeastern), Toaripi, Yaqui, Copainalá Zoque, Arabela, Aymara, Bora, Chayahuita, Cubeo, Embera (where the meanings are associated with different genders), Guaraní, Huambisa, Miskito, Piro, Ancash and Imbabura Quechua, Tsafiki, Yanomámi, Fijian, and Takia, and with 'spring' or 'well' in Buli, Wintu, Yuki, Cayapa, Kaingang, and Maxakalí (by the analyzable term *kōnāgkox /kōnā'āg-kox/* 'water-hole'). Given the possible variation in the size of lakes, from very small to very large, and the fluid boundaries between a lake and a pool down to a small pond, the meaning 'puddle' is also colexified in 24 of the sampled languages, namely Efik, Baruya, Buin, Kyaka, Rotokas, Kosarek Yale, Abzakh Adyghe, Cheyenne, Haida, Lesser Antillean Creole French (also colexifying 'mast, pole'), Tuscarora, Copainalá Zoque, Bora, Cashinahua, Guaraní, Ancash Quechua, Hawaiian, Lenkakel, Rotuman (by the analyzable terms mentioned above), Sedang, and Yay. Cubeo *macajitab* is, however, peculiar, apparently consisting of *maca* 'faeces' and *jitab* 'puddle.'

In two Polynesian languages spoken on small islands, Hawaiian and Fijian, the meaning 'inland' or 'interior' is also colexified; in Hawaiian, there is an optional complex term *loko wai* 'interior/lake water.'

Other associations are also found cross-linguistically. Buli *biung* is primarily used to refer to a “watering hole that dries up quickly” and rarely assumes the meaning ‘lake, pond.’ Hausa *tabki* ~ *tafki* also refers particularly to “water in a borrow-pit” (it is also an “exclamation of astonishment at bigness”) and *bingi* also to a “‘rough-coated’ fowl” as well as “[a]ny large donkey” inter alia. Swahili *ziwa* also means ‘breast,’ Kwoma *naba* is also used to refer to a ‘waterway’ and ‘canal,’ Kyaka colexifies ‘lake’ with ‘shallow water,’ and Ngaanyatjarra *murrkunggu*, denoting a ‘salt lake’ specifically, is by metonymy also used for ‘salty soil,’ while *parntu* can also refer to ‘salt’ itself. Badaga *aḷla* ~ *haḷla* also may refer to a “bottomland, lowest spot, depression” and ‘floodwater.’ Khalkha *nayur* can also refer to a “dry lake bed.” Welsh *llyn* is also used with the meaning ‘liquor, drink’ and *llwch*, an obsolete term for ‘lake,’ also means ‘dust, powder.’ Kolyma Yukaghir *jalyil* is also the name for the tambourine of a Shaman, and *ńorol* colexifies ‘moss.’ Kashaya colexifies ‘slick,’ Lakhota ‘I went,’ the reduplication base of Lake Miwok *pólpol* means ‘float, flood’ (although *pólpol* may be a loanword from Cache Creek Patwin as a whole), and Wintu *ḷul*, also denoting a ‘water hole,’ may be etymologically connected to a word for ‘to bubble’ (there is also the term *sa’wal* for a mountain spring with mythological significance). Cavineña *bei* is also the name of the ‘anteater,’ Hawaiian *loko* also means ‘in, inside’ and may refer to the ‘internal organs’ inter alia, while *moana* is, alongside its usage to denote both ‘lake’ and ‘ocean,’ also applied to name a “campground, consultation place for chiefs” (the common denominator apparently being that both are an expanse). Bwe Karen *nò* also denotes a ‘moat’ and a “written musical note” (perhaps, as the source suggests, due to English influence), while Lenakel *nisiu* also denotes a particular lake, namely Lake Siwi on Tanna Island, Vanuatu. Mandarin *hu2* also means “bottle-gourd, flask, teapot” (in both meanings going back to Early Middle Chinese *yo*, Pulleyblank 1991: 126), Sedang colexifies ‘naked,’ and by another term ‘ripe, red,’ and Lesser Antillean Creole French *lak* also may refer figuratively to a “slackening of control.”

### 35. *The Lightning*

Representation: 93 %

Motivated: 39.9%

Therof Analyzable: 22.0%

Thereof Colexifying: 18.4%

Thereof by Contiguity: 12.6%

Thereof by Similarity: 22.0%

Recurrent associated meanings: thunder, gleam/lighten/shine, electricity, light, spark, spirit/god, rain, fire, thunderstorm, sky, eye, telegraph, flame, arrow, blink, sword

Metaphorical denominations for ‘lightning’ (or ‘thunderbolt’) are common in the world’s languages, and some of them have a remarkable areal distribution (see § 6.4.3.). In Eurasian languages (represented by Khalkha and Kildin Saami in the sample), complex terms for ‘lightning’ using the source concepts ‘thunder’ and ‘arrow’ are found (*ajungya jin sumu* ‘thunder GEN arrow’ and *tīr’m’es’-ḡull* ‘thunder-arrow’ respectively). Another pattern predominantly found in Eastern Eurasia (Ket and Kolyma Yukaghir), but also in languages of the American Northwest (Central Yup’ik and Kashaya) is a lexico-semantic association

with 'fire.' Ket has *ekkinna boʔk*, analyzable as /*ekkin-na boʔk*/ 'thunder-3POSS.AN.PL fire,' Kolyma Yukaghir *jedun-ločil* 'thunder-fire,' Kashaya *maʔkala ʔoho* 'thunder fire,' and Central Yup'ik *kenerpallak* /*keneq-pallag*/ 'fire-intensively;' there are semianalyzable terms where 'fire' is the meaning of the identifiable constituent in Chukchi and Kildin Saami (the association is also weakly attested in Indo-European, Buck 1949: 57). Areality with respect to the meaning 'lightning' is also found in languages of Southeast Asia: the source concept 'blink' is encountered in Manange and White Hmong (Manange *tipli-1kʰa* 'blink-come' and White Hmong *xob laim* 'xob blink,' see Bauer 1992 for further discussion). White Hmong, by way of having a term referring to a spirit named *Xob*, also participates in another putative Southeast Asian pattern, which is precisely characterized by making reference to a spirit: Yay has *pya³ θaw⁵* 'pya³ throw' and *pya³ ta⁴ yaay⁶* 'pya³ pull/draw sword,' *pya³* being the spirit causing thunder and lightning (note in this context also Japanese *kami-nar-i* 'god-sound-NR,' as well as the fact that Burarra *andarrbaykarda* is also the name of a lightning spirit, and compare also the association between 'lightning' and 'Thor's hammer' in Old Norse and with 'Perkun's hammer' in Lettic). In turn, the association with 'sword' in Yay is also likely to be part of an areal pattern of Southeast Asia, the evidence being that Sedang has *cháng tōro* 'sword/men's tool thunder.' One of the source concepts in two languages of New Guinea, Meyah and One, is 'cloud;' the Meyah term is *mocgój efésa* 'fog/cloud flash,' the One term is semianalyzable only. Two languages, Tsafiki and Tetun, have an association between 'lightning' and 'flame' (*cunta pinda* 'thunder lightning/flame' and *rai-lakan* 'land flame' respectively; note also Yaqui *yuku be'oktia* /*yuku be'ok-t(e)-ia*/ 'rain-lightning-INTR-NMLZ' which also means "to take out the tongue several times" and the semantic connection between 'tongue' and 'flame' reported in section 22; Yaqui also has the term *yuku jima-ri* 'rain throw-RES').

Unsurprisingly, 'lightning' is also associated with meanings such as 'to gleam,' 'to lighten,' 'to shine,' or the like (common also in Indo-European, Buck 1949: 56-57). Hausa *walk'iya* also means 'glossiness' alongside 'lightning,' Toaripi *kevaro* also means, often reduplicated, "the shine, gleam on the leaves of plants and trees, or on a person's skin; the flash or sparkle of anything bright." Associations like these are also found by colexification in Abzakh Adyghe, Arabela, Bororo, and Lenakel, while Efik and Toba have derived terms, Sora *'kila:igum* is analyzable as /*'kila:i-gum*/ 'shine.brilliantly/dazzle-rain' (there is another semianalyzable term with a constituent meaning 'rain' in Sora), Carib *kapekape* is reduplicated from *kape* "smoothness, gleam," Guaraní *aravera* is analyzable as /*ára-vera*/ 'sky-brilliance' (Guaraní also has *aratiri* /*ára-tiri*/ 'sky-crack;' for the association with 'sky,' compare also Kwoma *neer hopo* 'sky snake' and Hani *aoq-miaovq miaovq* sky-burn RED), and Tetun *rai-nabilan* as 'land-shine.' There is a semianalyzable term where the identifiable constituent means 'to shine' in Hani. Another association that is akin to that just discussed is that with 'light,' occurring in six sampled languages. In Chukchi, *jəŋqeryen* (containing *jəŋ* 'fire') means both 'light of fire' and 'lightning,' and in Rendille 'lightning' is colexified with 'flash(es) of light' generally. Lenakel *nasiapumelaan* is derived from *asiapumel* 'to lighten, flash,' which in turn contains *asia* "[m]ake a light, make a torch (from coconut fronds)." If the association is by analyzability of the lexical type, there is variation in the semantics of the second element: in Berik and San Lucas Quiaviní Zapotec, it is



‘thunder’ (*iris naf* ‘thunder light’ and *x:cha’ x:tèe’ bziu’* ‘light GEN thunder’ respectively); in Maxakalí, it is ‘rain’ (*tex yānām* ‘rain light’). In fact, as already seen in the Yaqui example cited above, ‘rain’ is also a meaning that is frequently associated with ‘lightning.’ Alongside Sora, Yaqui, and Maxakalí, which were already discussed, the pattern is also found in Mbum, where ‘lightning’ is either *sāk à mbām* ‘tear GEN rain’ or *máà-mbām* ‘mother-rain.’ An association with the ‘eye’ is found in three languages of North America, Ineseño Chumash, Nez Perce, and Tuscarora, although it is doubtful whether they represent the same conceptualization strategy. In Ineseño Chumash, the ‘lightning’ is called *štix a soxk’on* ‘the eye of the thundercloud.’ In contrast, in Nez Perce, it is called *taqasaʔyóxoʔt /teqe-se-ʔiyóxoʔ-t/* ‘suddenly-eye-watch,’ suggesting that this term is not metaphorical in nature, but rather refers to the sudden perception of light by the human eye. The same is true of Tuscarora *newatkahréhnari:ks*, containing the roots *-kah(r)-* ‘eye’ and *-rik-* ‘bite.’ In Lakota, a literal translation of the word for ‘lightning,’ *wakíyatuwápi*, would be ‘the thunderbirds are looking’ (compare the notion of the thunderbird in North America mentioned in § 6.4.3.15.5.). Further complex terms for which the internal structure is not entirely clear include Kiowa’s *bəuəbɪhɛiɪgɪn* which appears to contain the root *bəuə-* ‘transparent,’ Sko’s *hénghèng* (putative reduplication base: *hèng* ‘fart’) and Bororo’s *baigabe* (compare *baiga* ‘Bororo bow’ and *be* ‘excrement’).

An interesting conceptualization of the meaning ‘lightning’ is found in Haida, involving verbal classifiers and the word for the colour ‘red’: *sri q’asda* contains *sri* ‘to be red’ and the classifier *q’a* for loud sounds (or possibly for large twodimensional surfaces) and *srid raaʔuhlda* the classifier *raa* for flashing light and *ʔuhlda* ‘to blink eyes’ (for which compare the association with ‘eye’ in other North American languages discussed above; note also that Comanche *ekakwitseʔe* may contain *eka-* ‘red’).

As already seen in various examples, complex terms for ‘lightning’ unsurprisingly frequently are made up in part of terms for ‘thunder.’ Other complex terms with that structure other than those already mentioned are Japanese *raku-rai* ‘fall-thunder,’ Carrier *t̥t̥ni elkreš* ‘thunder emit.flashes,’ and Cayapa *cuidya pi’queno* ‘thunder little’ (obviously of relatively recent vintage due to the presence of the loan from Spanish). However, ‘thunder’ and ‘lightning’ are also frequently associated by colexification, which is the case in 21 sampled languages: Buli, Ngambay, Rendille, Yoruba, Kwoma, Tasmanian (dubiously), Abzakh Adyghe, Japanese, Itzaj, Xicotepec de Juárez Totonac, Yana, Yaqui, Abipón, Arabela, Aymara, Hupda, Jarawara, Tehuelche, Bislama, and Bwe Karen. In Itzaj, the meaning ‘lightning’ can be singled out by the complex term *jatz’ chaak* ‘whip thunder/lightning,’ and in Takia by the verbal terms *weil i-fni* ‘thunder/lightning 3SG-hit’ and *weil i-raklawi da* ‘thunder/lightning 3SG-wink IPFV’ (for this association compare Efik *ekepkep*, derived from *kep* meaning ‘to corruscate, flash, lighten,’ but also ‘to wink’). There is a semianalyzable term with an identifiable constituent with the meaning ‘thunder’ in San Lucas Quiavini Zapotec, and in three further languages, Ngambay, Burarra and Tetun, ‘thunderstorm’ or ‘electrical storm’ is additionally colexified.

By colexification, associations with ‘spark’ are found in six sampled languages, namely Abzakh Adyghe, Badaga, Abipón, Bora, Ancash Quechua and Lesser Antillean Creole French. Yanomámi colexifies ‘lightning’ with ‘electric spark’ particularly, and indeed, a

relatively common pattern of semantic extension is to use the term for 'lightning' also for 'electricity.' This is found in Swahili, Basque, Khalkha, Nez Perce, Pawnee, Hupda, Hawaiian, and Samoan (the relevant Khalkha and Nez Perce terms colexify also 'telegraph').

Other associations include: the general meaning of Buli *ngmoruk* is 'rain,' and it rarely refers to 'thunder' and 'lightning.' Ngambay *ndàngè* also means 'to scold someone' and 'hurry,' and *tèl* also means 'suffering, pain' and 'announcement.' Muna has *bhibhito*, containing *bhibhi* 'to quiver, shake,' Nunggubuyu *-marawadja-*, containing *-w<sub>2</sub>adja-* 'to hit, kill,' and for Rotokas *parakau*, compare *paraka* 'wide, spread out,' 'width' and the classifier for narrow objects *ua*. Colexification with 'hailstone' is found in Basque, and Japanese has *inazuma*, analyzable as /ine-tsuma/ 'rice-spouse.' In Khalkha, *vacir ~ vcir ~ ocir* is also the name of a sacred instrument in Buddhist ceremonies, Cheyenne *vovó'ho'kâsé'há* contains *vovó'h* 'spotted white' and also means 'to flash repeatedly,' Itzaj *lemlem ~ lenlem* appears to be reduplicated from *lem* 'calm,' while in Jarawara, the same term is uniquely (with respect to the sample) also used for the 'sun' (and by extension also 'clock,' see discussion in section 79) and both 'lightning' and 'thunder.' Wichí has *pelhach'e /pelhay-ch'e/* 'storm-stick,' Bislama *laetning* also means 'in excess' and 'very quickly,' and Fijian *liva* also 'lift, lever' (in this sense due to borrowing from English). Kapingamarangi *ila* is also the term for a 'mole on the skin' or a 'maggot,' Rotuman *mere* also means 'to criticize, find fault with,' and colexification with 'wheel' in Samoan is almost certainly secondary due to English influence - the relevant term is *uila*.

Finally, it should be mentioned that in many languages, the semantics of 'lightning' is primarily or exclusively encoded verbally and the noun being derived from it, due to the temporal instability of the concept and its essentially event-like nature. This is for instance the case in Khoekhoe (*napa-b ~ tapa-b ~ lapa-b* 'to.strike.as.lightning-3SG.MASC') and in Laz, where the verb *divalai* is used to refer to the event of lightning.

### 36. The Meteoroid/Shooting Star

Representation: 48%

Motivated: 61.5%

Thereof Analyzable: 55.9%

Thereof Colexifying: 5.6%

Thereof by Contiguity: 5.3%

Thereof by Similarity: 54.3%

Recurrent associated meanings: star, tail, faeces, fall, fly, fire, run, spark

Terms for the meteoroid (for which 'comet' was accepted as a proxy if no other terms could be retrieved from the sources) are most frequently of the lexical type, one of the constituents being 'star' (or very rarely, a term with a fire-related meanings such as 'spark,' see below for further possibilities) and the other standing in a metaphorical relation with the target concept 'meteoroid.' In some languages, the second constituent is verbal and means either 'to fall,' 'to fly,' or less commonly 'to run.' For instance, One has *leila fanta yolu* 'star ?? fall,' a pattern also found in Khoekhoe, Toaripi, Cheyenne, Yuki, Bislama, and Tetun, Yuki has *mančipq:se ʔiʔ-ik* 'star fly-??,' a pattern also found in Khalkha, Hawaiian and Samoan, and Chickasaw *fochik malili-* 'star run-NMLZ,' a pattern also found in Abzakh Adyghe and Kapingamarangi. Even clearer metaphorical denominations are found when the

second constituent is nominal in nature: recurrent patterns of this kind include a metaphorical comparison with ‘tail,’ as in Yoruba *iràwòonírù* /iràwò-oní-irù/ ‘star-owner-tail’ and also in Ineseño Chumash and Rama (where an additional element meaning ‘sending’ is present), or, a little more commonly, ‘faeces,’ as in Highland Chontal *tifay galxamna* ‘faeces-star,’ and also in Haida, Central Yup’ik (where ‘meteor’ is colexified with ‘puffball,’ since “meteors are traditionally said to turn into puffballs when they land”), Toba (where the relevant term colexifies ‘meteoroid’ with ‘mushroom,’ Toba also has *to’olloxoi*, apparently derived from *to’olloxoi* ‘mushroom’ by means of the masculine suffix -c; for the association between ‘mushroom’ and ‘faeces,’ see section 41 and § 6.2.3.3.), Sedang, and Tetun (where an additional element meaning ‘hurl’ is present). Further complex terms involving ‘star’ are Basque *izar koloka* ‘star loose,’ Khalkha *suyunay odu(n)* ‘column.of.smoke star,’ Carrier *sem-thelšek* ‘star-uses.to.go.off,’ Upper Chehalis *scak<sup>w</sup>útwaln tat łačis*, containing *łačis* ‘star’ and the reciprocal marker *-twal*, presumably among other morphemes, Kashaya *q<sup>h</sup>a<sup>m</sup>oš šu<sup>h</sup>u<sup>h</sup>ta<sup>h</sup>du*, analyzable as /q<sup>h</sup>a<sup>m</sup>oš šu<sup>h</sup>u<sup>h</sup>ci<sup>o</sup>d-w/ ‘star by.pulling-pieces.come.off.bigger.object-DUR-ABS,’ Bororo *ikuieje ukigareu*, containing *ikuieje* ‘star’ and *kigareu* ‘adorned,’ Miskito *slilma dakwi ba*, consisting of *slilma* ‘star,’ *dakwaia* ‘break’ and the demonstrative element *ba*, Piro *katagiri psojite* ‘star fragment,’ and Fijian *kalokalo cavu* ‘star eradicate’ (with *cavu* also having other meanings, among them ‘adorned, highly decorated’). Moreover, Rotuman *hef sāl’ak mala* contains *hefu* ‘star’ and *mala* meaning ‘red hot’ but also denoting a red kind of belt worn by high chiefs, and a further San Mateo del Mar Huave term for the ‘meteoroid’ is *nandaab ocas* ‘burnt star.’ Colexification with ‘star’ itself (as well as ‘planet’) is encountered in Toaripi and Sora, while in Khoekhoe, the same term suffixed with different nominal designants yields the meanings ‘star’ and ‘comet’ respectively; in both cases, the relevant terms are derived from a verb meaning ‘to blink, twinkle.’ In Bislama, *sta* is glossed as “any heavenly body (e.g. moon, star, meteorite).” There are semianalyzable terms where the identifiable constituent is ‘star’ in Mbum, Sko, Biloxi, and Chayahuita. None of the abovementioned patterns has a clear areal hotspot of occurrence, rather, each one recurs in many different areas of the world.

There are also variants of some of the associations in which, rather than ‘star,’ some other meaning figures. Guaraní has *jagua-veve* ‘dog-fly,’ and Tuscarora *nekačis-nahkwá?ne?* contains the roots *-či’sn-* ‘spark’ and *-ne’ne-* ‘fly’ (for which compare Hani *miqzaq miqseil* ~ *aqzaq miqseil* ‘fire spark;’ *miqseil* is also “a woman who is dissatisfied with her marriage”). Variants of the denomination via ‘tail’ are San Mateo del Mar Huave *miwiül ix* ‘tail rock.iguana’ and Dongolese Nubian *káǵ-n-ǵu* ‘horse-GEN-tail’ (indeed, both terms can also be interpreted literally and refer to an iguana and horse tail respectively). Badaga is the only language in the sample where this association is realized by colexification (also colexifying “slender means” and “anything meagre”). Lake Miwok *wiküwiki* is a reduplicated version of *wiki* ‘fire,’ while Itzaj colexifies ‘shooting star’ with ‘fire’ and other meanings directly. Guaraní *yvarata* is analyzable as /yva-tata/ ‘fruit-fire’ (there is also a semianalyzable term where the identifiable constituent is ‘sky’).

Another strategy is found in Middle-Eastern Tasmanian, where *pökarit’ē* appears to contain a constituent meaning ‘ghost’ (compare the annotation for Kwoma *maway*: “Shooting stars are thought to be the souls of outstanding warriors who have died. At the

second, or final, burial of such a man his soul, in anthropomorphic form, flies through the air holding a burning coconut frond torch in its hand, which is the light seen in the sky, towards one of the many large lagoons that surround the Washkuk hills”). Great Andamanese *chàugalalachōinga* is somewhat similar in that it makes reference to the ‘spirit’ (*chàugala*); *chōinga* means ‘light of torch.’

Other unique denominations are found in Efik, where *inōntanta’fioñ* contains *i’nō* ‘thief’ and *ika* is derived from *ka* ‘to go.’ Sahu has *deterē’e ma ngi’di* ‘tooth POSS thunder,’ and Welsh *maen mell* ‘stone lightning’ alongside *awyrfaen* /*awyr-maen*/ ‘air/sky-stone.’ Basque *meteor* is also glossed as ‘atmospheric phenomenon.’ A literal translation of Kiliwa *xsmii??aawpmaay* is ‘Xsmii’s fiery urine’ (Xsmii is the name of a constellation; compare the associations with ‘faeces’ discussed above). Wintu *nomleyna*’s is literally ‘going west.’ For Abipón *neiak*, compare *eiagaiag-* ~ *-eiagaiak-* ‘walk, travel.’ Miskito has *imyula mabra* ‘lightning egg,’ while Piro *gijrukachri* appears to contain *gijru* ‘shaft, handle.’ Yanomámi *thoru wakē* consists of *thoru*, the name of a plant species, and *wakē* ‘red,’ and is at the same time the name of the fire spirit. Patterns of colexification include that with ‘spear’ and “standing upright on the hands” in Hausa, “a flaming arrow, as a signal, or for setting fire to enemy villages” in Buin, that with ‘angel’ in Rotokas (the term, *purapurapato*, is likely derived from the verb *pura-* ‘to contract something, say, make, do, create’ and the derivational element *-pa*), that with ‘Johnny Jump ups’ (a species of violets) and ‘flowers’ generally in Wintu, that with ‘stick, crop’ in Aguaruna, and that with ‘thunder without clouds’ in Arabela.

### 37. The Milky Way

Representation: 38%

Motivated: 70.3%

Thereof Analyzable 65.3%

Thereof Colexifying: 5.9%

Thereof by Contiguity: 1.8%

Thereof by Similarity: 66.5%

Recurrent associated meanings: trail/road/street, star, sky, river, eel, ghost, gird, back-bone, tapir

For terms for the Milky Way (or ‘galaxy,’ which was accepted as a proxy), it is the rule rather than the exception to be motivated. Apparent exceptions are found in Katcha, Khoekhoe, San Lucas Quiaviní Zapotec, Nunggubuyu, Great Andamanese and Kapingamarangi, for which no analyzability can be safely diagnosed on the basis of the source and no further meanings are stated. Lexico-semantic associations are often culture-specific and bear mythological connections. For instance, Rendille *íntí warába abártiís kájiité* is literally translated as ‘the place where the hyena dragged his mother’ and comes from a Rendille children’s story of the same name, and in Ket the name for the Milky Way is *Albakaŋ*, analyzable as /*alba-kàŋ*/ ‘Alba-way,’ Alba being a Ket mythological hero (Andrej Nefedov p.c.). In spite of the highly language-specific associations for the Milky Way, there are nevertheless some general naming tendencies, and the Ket example already provides an example of this: in fact, a recurring denominational pattern for the Milky Way found also in many well-known European languages is that involving various travel paths,

such as 'trail,' 'road,' 'street,' and also in a few cases 'river,' with the second element usually language-specific. Thus, Muna has *sala waghua*, containing *sala* 'path, trail' and presumably *ghua* 'gray hair,' Blackfoot has *makóyoohsokoyi* /makóyi-mohsokó-yi/ 'wolf-road-INAN.SG,' Upper Chehalis *sxai-yoxoʔ* 'trail-poor.people,' Cheyenne *Áméó'o* is analyzable as /ame-meo'o/ 'pemmican-way' and *séotsé-méó'o* perhaps as 'corpse-road' (the source notes difficulties with the precise translation), Chickasaw *Ofi' Tohbi' Ihina' ~ Ofi' Tohbi'hina'* as /ofi' tohbi-' im-hina'/ 'dog be.white/be.pale-NMLZ DAT-road,' Itzaj has *sák-b'ej* 'white/gray-trail/road' (colexifying 'highway'), Lakhota *wanáǵithačhąku* /wanáǵi-tha-čhąkú/ 'ghost-??-road,' Nez Perce *čewčé'w-nim ʔiskit* is 'ghost-POSS trail,' Wintu *qanal yemer* 'be.open.wide road' and *qaqal yemer* 'oblivion road.' Chayahuita *chimirin ira* is 'death way,' Kaingang *krīg japyry* 'star way' (there is a further semianalyzable term with *krīg* that is also the name of a particular star), Miskito *swara bila* 'eel way' (*bila* also has other meanings alongside 'way'), Rama *núnik kás aríra*, a literal translation of which is 'cloud street' (although *aríra* is glossed as 'string, fold' on its own), Wayampi *tapiʔi-la-pe* is 'tapir-of-way' (a further variant being *tapiʔi-lipə* 'tapir-track,' compare Guaraní *mborevirape*, also containing *mborevi* 'tapir' and *tape* 'way'), Fijian *sala-ni-cagi* 'road-POSS-wind/air,' Sedang *tróang hólóng* 'road star' (alongside the alternative term *holóng nhéng* 'star mirror').

As the examples show, the Milky Way is often associated also with otherworldly phenomena, such as 'ghosts' in Lakhota and Nez Perce and 'death' in Chayahuita, but by no means necessarily so. Nor is it the case that the association with travel paths is always by morphologically complex terms, Hausa is an example of a language with colexification (here, also 'channel,' 'intermediary,' and 'beehive' are colexified inter alia). A further case of colexification is perhaps Tasmanian (Plomley 1976: 408), and Muna has a semianalyzable term.

Yir Yoront and Ancash Quechua colexify 'Milky Way' with 'river' directly (in Ancash Quechua, there is the optional complex term *paqas mayu* 'night river'), Chukchi has *cəyɛj-weem* 'sandy-river' and Tetun *mota-klakatak* 'river-reflection/image' and *mota-leten* 'river-top/summit.' In Khalkha, an association with a 'gird' is found by a morphologically complex term (*oyturyui jin byse* 'sky GEN girdle' alongside *tngrɪ jin ojudal* 'sky GEN seam'), and one of the glosses of Hawaiian *kau* is 'to gird.' The association with 'milk' (the tertium comparationis obviously being the whiteness) is restricted to languages of Europe in the sample: Basque has *esne-bide* 'milk-way' (as well as the alternative term *santiago-bide* 'Santiago-way') and Greek a synchronically semianalyzable term. However, there are also languages outside Europe in which associations with whiteness seem to occur. For instance, Sahu has *ka'e ma geolo* /ka'e ma ge'olo/ 'drink/palm.wine POSS foam,' Kiliwa *?+ma?i=ny-?+phuuy* 'DN+sky=POSS-DN+smoke,' Bororo *kuiejedoge eerugudu*, containing the words for 'star' and 'ash' (there is also the alternative term *ikuiejedoge erugudu* containing *ikuieje* 'star' and *erugu* 'see'), San Mateo del Mar Huave *Minajndot Oleaj Micawüy Santiago*, containing *najndot* 'dust' and *cawüy* 'horse.' Note also Xicotepec de Juárez Totonac *i'xtej stā'ná' kaxtaj* (*kaxtaj*, 'lime') alongside the already mentioned Chickasaw and Itzaj terms.

In Hawaiian, the Milky Way is either *i'a*, the basic meaning of which is 'fish' (it can also refer to an 'eel' inter alia specifically, compare the Miskito association with 'eel' mentioned above), or *i'a-lele-i-aka* 'fish-jump-in-shadow' (and there is also the term *hōkū-*

*nohoaupuni* ‘star-rule’). As seen from the examples discussed so far, ‘star’ is as frequent as obvious meaning of constituents of complex terms. Alongside the languages already mentioned, an element with this meaning also features in Sora (*a-o:n-tu’i-an* ‘POSS-child-star-N.SFX’), Nuuchahnulth (*taatusaniut*, analyzable as /tatus-añut/ ‘star-along’), and Hani (*aqgeel caqkov*, presumably analyzable as ‘star ditch.around.house,’ although ‘star’ is more fully *aqgeel alsiq* or *aqgeelsiq*).

In two languages of North America, Ineseño Chumash and Kiowa, an association with the ‘back bone’ is found: *šnokok’ a saxiyi* ‘the spine of the night’ and *tɬ’-gɔumt’ɔu* ‘star-backbone’ respectively (in Ineseño Chumash, an alternative term is *’alsuyopoš* ‘the piñon gatherer’).<sup>6</sup> Tuscarora *yurɛhyáhuks*, containing the roots *-rɛhy-* ‘sky’ and *-huk-* ‘light up,’ can also refer to a ‘rainbow’ and the ‘Aurora Borealis,’ and the Yuki term *miɬ ʔon kaw* also betrays an association with ‘sky’: it is analyzable as /miɬ ʔon kaw/ ‘sky land light.’ Likewise, Hawaiian *lā-lani* might be analyzable as ‘sun/day-sky’ (though both putative constituents also have many other meanings).

Other associations include: in Welsh, the Milky Way bears the name of mythological castles: *caer Wydion* and *caer Arianrhod* (*caer* ‘castle’), and in Carrier it is called *yq-ke-tšilkrai* ‘sky-one-run.over.’ Pawnee *Rakiiraruhtuturuuhah* contains *uhak* ‘to pass in a line,’ and in Central Yup’ik *Tanglurallret* contains *tangluq* ‘snowshoe;’ it is so called because “because in legend it is the snowshoe trail of the raven.” Associations by colexification are, as noted above, much rarer, though not nonexistent: Hausa, by another term than that mentioned above, also colexifies “[t]he top of the head of a horse, between the ears” and “[t]he top of the occiput of human beings” inter alia. The Ngaanyatjarra terms *tjukal(pa)* and *yintirri* both also mean ‘tree ladder,’ Nunggubuyu *burumburun<sup>a</sup>* also “galaxy, cluster of stars,” and Khalkha *mecid* also denotes ‘monkeys’ or ‘apes’ (formally it is the plural of *meci(n)* ~ *beci(n)* ‘ape, monkey, ninth year in twelve-year cycle’). Hawaiian *kau* has, alongside ‘Milky Way’ and ‘fish, eel,’ also the meaning ‘to place, put’ and is the “name of a star in the northern sky that served as guide to mariners” inter alia, and there is also the term *lele-aka* ‘fly-shadow’ and a further term containing *lele* which colexifies ‘shark-sucker’ and ‘remora’ inter alia. For Kapingamarangi *ganiwa* compare *gani* ‘penis’? The relevant Lenakel term contains a constituent meaning ‘girls.’

### 38. *The Moon*

Representation: 100%

Motivated: 67.1%

Thereof Analyzable: 6.5%

Thereof Colexifying: 60.6%

Thereof by Contiguity: 52.5%

Thereof by Similarity: 5.2%

Recurrent associated meanings: month, sun, moonlight, season, night, albino, menstrual period, light, snail, bright, white, torch

<sup>6</sup> In the English-Kiowa section, the term is given as *tɬ’-gɔumt’ɔu* in which case it would be ‘sibling-backbone;’ this is likely a printing error, since under the lemma for *gɔumt’ɔu* the term for ‘Milky Way’ is stated to be *tɬ’-gɔumt’ɔu*.

Many of the analyzable terms found cross-linguistically are due to the fact that many languages of the Americas, but also some in Siberia, lack true lexical differentiation for 'sun' and 'moon,' and have a single term for the two luminaries (found in 17 sampled languages all in all: Blackfoot, Carrier, Upper Chehalis, Cheyenne, Chickasaw, Lakota, Nez Perce, Nuuchahnulth, Quileute, Tuscarora (according to an older source incorporated into the consulted source), Wappo, Wintu, Bora, Cayapa, Cubeo, Hupda, Macaguán, and possibly Tehuelche, where the relevant terms are very similar segmentally. In some of these languages, 'watch' and/or 'calendar' is also colexified due to the 'sun'-reading, see section 60 for fuller discussion, and sometimes there is the possibility of optional disambiguation. Another possibility is morphologically complex terms for the 'moon' containing the respective word for 'sun,' for instance Lake Miwok *káwul híi* 'night sun.' The latter situation is encountered also in two other sampled languages, Kolyma Yukaghir (*emin-pu:gu* 'night-sun') and Maxakalí (*māyōn-hex* 'sun-female/man's.sister,' for which compare Rama *tukán kumá* 'moon woman/female').

At times, however, complex terms with a different structure are encountered. The Abipón term *eergRaik*, which is interestingly also used for 'star,' is derived from *eerg*- 'to burn, sparkle,' and Central Yup'ik *unuggsuun* ~ *unugcuun* is analyzable as /unuk-cuun/ 'night-device.for' and generally means 'night-light,' but 'moon' in the dialect of Nelson island. Kapingamarangi *malama* is derived from *lama* meaning 'torch' (alongside 'dry coconut leaves') and indeed also means 'lantern,' a similar situation is encountered in Hawaiian (note also that one of the terms used for 'moon' in Miskito, *ingni*, also means 'light' and 'lamp,' and that the Yuki term is semianalyzable with the identifiable constituent meaning 'light' inter alia). Furthermore, Greek *selḗnē* is diachronically related to a verb meaning 'to shine,' and Wintu colexifies 'moon' with "something shining, bright, white, albino, rare" (Fijian *vula* colexifies 'partial albino,' the Hani term for the moon contains an element meaning 'white,' the Yuki term an element meaning "light, clear, clean, shine," and the association with 'brightness' and 'light' is also present in a group of Indo-European terms, Buck 1949: 53). Interestingly, there are also two sampled languages in which the word for 'month' is monomorphemic, and that for 'moon' (at least synchronically, since semantic shift is of course possible) secondary to it (Yay *roy<sup>5</sup> duan<sup>1</sup>* 'bright month' and San Mateo del Mar Huave *mūm caaw* 'mother month').

As far as colexification is concerned, a ubiquitous pattern is indeed that with 'month,' occurring in 98 languages of the sample and in all areas of the world (and reconstructed for the Proto-Indo-European level as well as present in many daughter languages, Buck 1949: 54), namely Buli, Efik, Hausa, Katcha, Khoekhoe, Koyraboro Senni, Mbum, Ngambay, Noni, Dongolese Nubian, Rendille, Swahili, Yoruba, Berik, Buin, Gurindji, Kwoma, Kyaka, Lavukaleve, Mali, Muna, Ngaanyatjarra, Nunggubuyu, Meyah, Rotokas, Toaripi, Sahu, Sko, Yir Yoront, Abzakh Adyghe, Badaga, Chukchi, Japanese, Khalkha, Laz, Kolyma Yukaghir, Acoma, Biloxi, Cahuilla, Upper Chehalis, Cheyenne, Chickasaw, Highland Chontal (colexifying also 'goddess'), Ineseño Chumash, Comanche, Haida, Kiliwa, Kiowa, Lake Miwok, Nez Perce, Nuuchahnulth, Oneida, Santiago Mexquititlan Otomí, Pawnee, Quileute, Xicotepec de Juárez Totonac, Wintu, Yaqui, Yuki, Central Yup'ik, San Lucas Quiaviní Zapotec (colexifying also 'comb' and "comb for pushing back woven material on a

loom”), Copainalá Zoque, Arabela, Aymara, Carib, Cashinahua, Cavineña, Chayahuita, Cubeo, Guaraní, Huambisa, Hupda, Jarawara, Kaingang, Lengua, Macaguán, Miskito, Piro, Ancash Quechua, Rama, Sáliba, Tehuelche, Toba (by the term *ca’agoxoic*, for which compare *ca’agaxaic* ‘always hurries/worries,’ the original gloss being ‘se apura siempre’), Tsafiki, Fijian, Great Andamanese, Hani, Hawaiian, Kapingamarangi, Bwe Karen (colexifying also ‘be green, blue’ inter alia), Lenakel, Malagasy, White Hmong, Rotuman, Samoan, Sedang, Takia, Tetun, and Bislama (obsoletely). This, together, with the colexification of ‘sun’ and ‘moon’ mentioned above, is chiefly responsible for the high percentage of terms for ‘moon’ with another colexified meaning. Buli, Bislama, Hawaiian, and Mandarin colexify ‘moon’ with ‘moonlight,’ and in Ngambay, Kyaka, and Samoan, alongside ‘month,’ the meaning ‘season’ is also colexified (in Kyaka also ‘weather’).

A pattern conspicuously found in two languages of Australia, Burarra and Gurindji, is that the word for ‘moon’ also is used to refer to a ‘snail’ (for Burarra, it is explicitly stated that this snail is white and moon-like in appearance). Similarly, in Nunggubuyu, the word for ‘moon’ also denotes ‘small white grubs’ and the ‘chambered nautilus’ (and in another Australian language, Yir Yoront, the relevant term is also glossed as “moonlike in colour and shape”); perhaps similarly, Fijian *vula* is also the name of a kind of bêche-de-mer, inter alia, and Rotuman *hulā* is also the name of a kind of sea slug – given that the terms seem either cognate or related by borrowing into Rotuman, it is even possible that this secondary meaning is inherited or contact-related. Itzaj and Bislama colexify ‘menstrual period,’ and Itzaj by another term also ‘bead’ and ‘crumble.’

Other associations include: Buli *chiik* is also the name of a “moon-amulet,” and traditionally the ‘soul’ inter alia. Efik *ö’fiöñ* also denotes a “circular figure in painting” and is presumably related to the name of two days of the week. Buin *EEKIO* is also used to refer to “a period of ten days.” Kyaka *kana* also means ‘rock, stone’ as well as ‘money, coin.’ Sko *ké* is also the name of a “post for hanging things” and means “catch, get, take, fetch” as a verb. Sahu colexifies “something round” generally. Toaripi *papare* contains *papa* ‘grandparent, ancestor,’ and Khalkha *sumija* also bears the meaning ‘Monday,’ which was however obsolete at the time the consulted source was written. In Kiowa, the relevant term *ḡá* can also refer to a ‘river’ and is the name of a game, Lake Miwok *koméenawa* is analyzable as /kóme-nawa/ ‘vagina-old.man,’ Nuuchahnulth colexifies ‘moon’ with ‘thimbleberry,’ and Central Yup’ik *tanqik* only means ‘moon’ in the dialect of Nunivak island, and denotes ‘brightness’ generally in other dialects. Cayapa colexifies ‘moon’ with ‘lowered, fallen,’ Tsafiki with ‘excrement,’ Yanomámi with a mushroom species, and Hawaiian with a “crescent-shaped fishhook” and an “eye of the snail at the end of its horn” inter alia. Rotuman *hulā* also means ‘seed, pip of fruit’ and other things, and Samoan colexifies ‘moon’ with ‘chief.’

### 39. *The Mountain*

Representation: 91%

Motivated: 53.7%

Thereof Analyzable: 11.4%

Thereof Colexifying: 42.7%

Thereof by Contiguity: 15.9%

Thereof by Similarity: 28.6%



Recurrent associated meanings: hill, forest, stone/rock, summit/peak,  
high/top, land/earth, scarp/slope, pile, mountain range, head, valley

Cross-linguistically, lexico-semantic associations on a global scale closely mirror those mentioned by Buck (1949: 23–24) for Indo-European. Analyzable terms are most frequently based on notions such as ‘high’ or ‘top,’ alongside colexification in Yoruba. This is the case in five sampled languages: Kaluli has *hen misiyo*: ‘land high.place/high.ground,’ Muna *ka-bhawo* ‘ABSTR-high,’ Miskito *il tara* ‘hill high,’ Rama *king-úp* ‘head/top-eye’ (there are further semianalyzable terms involving *úp*, for *taisup* compare *taik* ‘piece, thing, top, nose, penis, and for *kaisup* compare *kais* ‘tick?’), and Fijian *dela-ni-vanua* ‘top-POSS-land’ and *ulu-ni-vanua* ‘head/top-POSS-land’ (note also that White Hmong *roob* is probably a loanword from Chinese *chóng* ‘high, lofty’); Abzakh Adyghe has *q<sup>o</sup>ʔəʃhe* /(-)q<sup>o</sup>ʔ(e)-ʃhe/ ‘arm/branch/pointed.object-up.there.’ This is also the only clear instance of analyzable terms involving a body-part metaphor. Analogously to the analyzable term in Rama, ‘mountain’ and ‘head’ are colexified in Gurindji and Kaingang, where the relevant term also means ‘cockscorn.’

Terms for ‘mountain’ often also mean ‘forest’ by spatial contiguity, as already discussed in section 26. This is the case in Kwoma, Basque (colexifying ‘woodland’ more specifically), Khalkha, Laz, Nivkh, Ineseño Chumash, Aguaruna, Huambisa, and Yay (disregarding terms glossed as ‘selva’ and/or ‘monte’ in Spanish, since these also capable of referring to ‘wilderness’ more generally and thus the sense colexified is unclear due to this bias exerted by the metalanguage). Similarly, Sora colexifies ‘hill’ with ‘forest,’ Cubeo has *jocu-burumu* ‘wood-CLASS.VEGETATION’ (there is another semianalyzable term involving a classifier for vegetation that has not been cut down in Cubeo), and Embera uses the same term for ‘mountain’ and, associated with different genders, various kinds of plants as well as ‘plantation’ and ‘vegetation’ generally. Also by spatial contiguity, Buin and Tsafiki colexify ‘mountain’ with ‘valley.’

Alternatively, terms for ‘mountain’ may contain constituent elements meaning ‘hill’ and an additional element frequently conveying an additional semantic component of big size or height: San Mateo del Mar Huave has *ti-tiüc* which appears to be analyzable as ‘AUG-hill,’ Cheyenne *ho'honáe-vose* ‘rocky-hill,’ Tuscarora *yunətherʔúy?*, containing *-nəth(e)r-* ‘hill’ and *-iyu-* ‘be great, be beautiful,’ and Miskito *il tara* ‘hill high.’ There are semianalyzable terms in Bororo and Lengua. As also noted by Buck (1949: 23) with regard to ‘hill,’ there are sometimes “fluctuating discriminations” between ‘mountain’ and ‘hill,’ and this is reflected insofar as that these meanings are also frequently colexified. This is the case in Buli, Ngambay, Dongolese Nubian, Yoruba, Burarra, Gurindji, Kwoma, Kyaka, Muna, Ngaanyatjarara, Tasmanian (Northerastern, Middle-Eastern, and Southeastern), Toaripi, Sentani, Yir Yoront, Abzakh Adyghe, Badaga (where there is a verb of the same phonological structure meaning ‘to cut’), Ket, Upper Chehalis, Cheyenne, Chickasaw, Highland Chontal, Itzaj, Kiliwa, Lake Miwok, Lesser Antillean Creole French, Santiago Mexquititlan Otomí, Pawnee, Tuscarora, Wappo, Yaqui, Yuki, San Lucas Quiaviní Zapotec, Bora, Bororo, Embera (where the meanings are associated with different genders), Hupda (also colexifying ‘cliff’), Jarawara, Lengua (colexifying ‘large hill’ specifically), Macaguán,

Maxakalí (colexifying ‘steep hill’ specifically), Ancash Quechua, Rama, Tsafiki, Wayampi (dialectally), Yanomámi, Bislama, Hawaiian (colexifying ‘high hill’ specifically), Kapingamarangi, Malagasy, Rotuman (colexifying also “mound of earth or sand” inter alia), Samoan, Takia, and Tetun; note also that for Toaripi *raepa*, glossed as ‘hill, mountain,’ the lexicographer remarks that “[f]or the latter meaning the adj[ectives] rovaea (=big) or koa (=high) is often added.”

As already seen in some terms, complex terms for ‘mountain’ at times also feature a constituent with the meaning ‘land’ or ‘earth’ generally. Further instances of this are found in Efik (*akamba obüt* ‘great land/earth’), Kosarek Yale (*mok-soo* ‘place-earth,’ this term also means ‘world’ and ‘land’ by itself), Guaraní (*yvy-ty* ‘earth-pile’), Wayampi (*iwi-ti* ‘ground-big’), and possibly Cheyenne (*o’omenó* may contain *-o’ome* ‘region’ and *-nó* ‘place’). There is a further semianalyzable term where the identifiable constituent is ‘land’ in Kaluli.

Buli, Abzakh Adyghe, Jarawara, and Tetun colexify ‘mountain’ with ‘scarp, slope,’ and Ngaanyatjarra, Yaqui, and Mandarin with ‘mountain range.’ There is, as also noted by Buck (1949) for Indo-European, a recurrent association between the meanings ‘mountain’ and ‘stone, rock,’ in the languages of the sample in Buli, Ngambay, Ngaanyatjarra (these two languages also colexify ‘pebble’), Kolyma Yukaghir, Biloxi (where the colexified meaning is “round-topped hill” more specifically, the relevant term may contain an element meaning ‘sharp’), Cahuilla, Oneida (also colexifying ‘outcropping’), Hupda (colexifying also ‘sky’), and Hawaiian (where “kind of hard stone from which adzes were made” is in addition colexified). White Hmong has *pob-tsua* ‘ball-rock.mass,’ and there is a semianalyzable term featuring a constituent meaning ‘stone’ in Copainalá Zoque. In Kyaka, *kyau* also means ‘pile’ and ‘swelling,’ Highland Chontal *tijuala* also means ‘pile,’ and Lesser Antillean Creole French *mòn* may also figuratively refer to ‘a large heap’ (compare Lenakel *touar* ‘mountain’ and *tou* ‘to heap up food, yam heap?’), while Kiowa *koup* may also refer to a ‘knob’ (again, compare Buck 1949: 23); there is also a verb of the same form meaning ‘to lay several.’ At times, the relevant terms also denote the ‘summit’ of a mountain or ‘peak’ more specifically, a case of meronymy. This is found in Baruya, Burarra (“top of rocky outcrop” more specifically), Abzakh Adyghe, Badaga, Khalkha (‘flat mountain top’ more specifically), Wintu, and Ancash Quechua; note also Abzakh Adyghe *ʔa-šhe* ‘summit-up.there.’

Other associations include: Dongolese Nubian *ǧébel* also means ‘desert’ (a meaning which is also colexified in Sudanian Arabic, from which the term is borrowed). Buin *kumpa* is also the name of an ‘edible mountain fern’ and a male name (the relevant term is, however, a poetic epithet for ‘mountains’ only), Rendille *hál* also means ‘pack camel’ and also denotes a particular mountain, Kwoma *kwow* also means ‘netbag’ and ‘womb,’ and One *ala nala* ~ *ala nela* appears to contain *nala* ~ *nela* ‘tooth’ (there is a variant *ala palla*, for which compare *palla* ‘body hair’). Sko *pì* also means ‘full,’ and Abzakh Adyghe *bǧə* also ‘waist’ and ‘backrest.’ Adjectivally, Badaga *male* also means “puffed up, haughty, self-important, towering above,” and Basque colexifies ‘mountain’ with “country, wild country.” Bezhta has an unusual pattern of colexification in that *mā* is also the default term for the ‘nose’ (note that Abzakh Adyghe *čape* might contain *p(e)*, meaning ‘nose, beak, point’ inter alia).

Khalkha colexifies 'mountain' inter alia with 'plateau' as well as "shelf, hanging rack," Ineseño Chumash with 'north,' Itzaj with 'pyramid,' Santiago Mexquititlan Otomí with 'dark,' and Tuscarora with 'billow.' For Lesser Antillean Creole French *montani*, compare *montan* 'rising, ascending,' and for Yana *ziigal(la) ~ ziigal(xi)* (which are dialectal variants), compare *zii-* 'smoky, foggy.' Copainalá Zoque colexifies 'mountain' with 'field,' and the Abipón term *lkaagRanRat* appears to be analyzable as *l-kaag-Ran-Rat* 'POSS.INDEF/3SG-split-CAUS-CAUS' (original Spanish gloss of the root *aag-* ~ *-aage* ~ *-aak* ~ *-kaak* ~ *-kaag* is 'morder, hender, partir, cavar'). Toba *qasoxonaxa* is also the name of a mighty being causing lightning, and is also used to refer to the 'elephant.' Hawaiian has *kuahiwi*, containing *hiwi* 'sharp ridge of mountain' and *kua*, meaning 'back, rear' inter alia. Kapingamarangi *gono duu* is analyzable as 'form/surface/hue stand/stop/belt.' Bwe Karen colexifies 'skin, shell' inter alia, and Lenakel "elevated place, village, township." Rotuman *solo*, as a verb, also means 'for the sun to sink,' Sedang colexifies 'mountain' with 'stump,' and Tetun with "country(-side)."

#### 40. *The Mushroom*

Representation: 69%

Motivated: 17.8%

Thereof Analyzable: 15.4%

Thereof Colexifying: 2.9%

Thereof by Contiguity: 1.0%

Thereof by Similarity: 15.7%

Recurrent associated meanings: ear, faeces, spirit/devil, horse

Terms for 'mushroom,' if they are motivated, have a very marked lexico-semantic profile: motivation is very often realized by analyzable terms, and the semantic relation underlying them is very often one of similarity. Although there is wide variation as to the specific semantic source concepts to be discussed below, there are two recurring strategies that are both remarkable. One conceptualization is based on 'faeces' (see also § 6.2.3.3.) and the other one is based on 'ear.' Thus, Rendille has *u'dú-yeyyah* 'moon-faeces,' San Mateo del Mar Huave *aonts potwit* 'excrete black.vulture,' Toba *huaqajñi l'-atec* 'star 3SG.POSS-excrement' (colexifying 'shooting star'), Hawaiian *kūkae-lio* 'excrement/dung-horse,' (this term also denotes grasses where horses are pastured. For the association with 'horse,' compare Toba *cainaton /caaño-naton/* 'horse-glans,' there are several variants of this term, and in one of them, the element denoting 'glans' is colexified with 'sombrero'), and Cashinahua resorts to direct colexification.

Carrier has *impiñ-dzo* 'pigeon-ear,' Central Yup'ik (Nunivak Island dialect) *tuunram ciutii* contains elements meaning 'spirit, devil' and 'ear,' Fijian has *daliga ni kalou* 'ear POSS spirit,' and Bislama *sora blong devel* 'ear POSS spirit.' Moreover, Aguaruna, Rotuman, and Samoan colexify 'ear' with (types of) 'mushroom' inter alia, and there is a semianalyzable term in Anggor; the Rotuman term *faliga* also means 'pectoral fin,' and there is also a complex term *faliga ne 'atua*, with *'atua* meaning 'ghost.'

The Samoan evidence, where the relevant term is said to denote several species of fungus, points to a potential problem: it is not possible to be sure that terms in the consulted sources really correspond to the life-form level and do not rather denote a specific

type of mushroom on the generic level. However, it seems unlikely that the cross-linguistically robust occurrence of the associations can be due only to such ambiguities due to dictionary information. For the area of Oceania, genealogical inheritance, areal spread, or a low diversity of biological diversity on the small islands of Oceania cannot be excluded as possible factors, but even then, the association is also found in areas of the world in which such factors seem unlikely. If the association with ‘ear’ is thus indeed a genuinely recurring phenomenon, what is the cause? Tree-growing fungi are quite widespread globally, and in some areas may even be the only type of mushroom found. Often, these fungi have a decidedly ear-like shape, as opposed to the typically cap-like appearance of soil-growing mushrooms. The Jew’s ear (*Auricularia auricula-judae*), particularly mentioned in the dictionary gloss for Samoan, is an instance of this.

There are, alongside those occurring in association with the metaphorical transfer of ‘ear,’ also other terms making reference to spirits: Wichí has *ahot-lhu* ‘spirit/soul-penis’ for ‘black mushroom,’ and Fijian furthermore *iviu ni tēvoro* ‘fan.palm POSS demon.’

There are further metaphorical patterns including associations with certain animals in languages of Africa: Hausa has *nama-n kaza* ‘meat-GEN hen’ (*kaza* also denotes an ungrateful person) and Kanuri *támbàl kókó-be* ‘drum frog-of’ (there is another term in Kanuri which appears to colexify ‘rainy season’); moreover, Wichí has *mawu-tonek* ‘fox-liver’ for ‘orange mushroom.’ Still other metaphor-based terms are Japanese *ki-no-ko* ‘tree-GEN-child,’ Kolyma Yukaghir *ara:n-paj* ‘naked-woman’ for a ‘mushroom growing on earth,’ Upper Chehalis *łúm=lwtx* ‘wrinkle/shrink=house/building/place.where.animal.lives’ and Kiliwa *phitnsmay* with the literal meaning ‘little lost fart’ (compare also Yay *rat<sup>3</sup> raap<sup>2</sup>* ‘mushroom’ with *taw<sup>3</sup> rat<sup>3</sup>* ‘to break wind?’). There is also an association with ‘hat’ or ‘cap’ in one language, Haida. There are several complex terms involving *dajing* ‘hat, cap’ one of them *kagann dajing*, with *kagann* meaning ‘mouse’ (“[m]ice were the physical form assumed by witches’ evil spirit”).

Other associations are: Efik *udip’ ek’pe* seems analyzable as ‘NMLZ-hide bunch’ (with *ek’pe* also meaning ‘panther, leopard’ inter alia). Ngambay *bbè* also means ‘quiver,’ and Yoruba colexifies ‘chief among persons’ and ‘queen of ants’ with ‘mushroom.’ Kaluli *ko:lo:* is also a “word to signal or point back to something just talked about or something just mentioned,” Sahu colexifies ‘mushroom’ with ‘rust,’ Basque *ziza* also means ‘lisping’ as well as ‘to excise tax,’ and Lesser Antillean Creole French *chanpiyon* also ‘champion’ (due to collapse of Fr. *champignon* and *champion*). Wintu *ʔaλ* also means to ‘look on, observe, watch’ inter alia, *qun* in the same language also means ‘mold’ and ‘blue.’ Central Yup’ik *palurutaq ~ paluqutaq* (Hooper Bay and Chevak dialect) also means ‘quonset hut’ and ‘turtle.’ Cubeo *chīchi* colexifies ‘scale,’ and Miskito *yula* is also used with the meanings ‘dog,’ ‘insects,’ ‘little animal,’ and ‘parasite’ inter alia, and *srapka*, another term, seems to contain *srap* ‘algae, moss.’ Piro colexifies ‘mushroom’ with ‘lichen,’ and Malagasy *hòlatra* also means ‘scar.’ In addition, the Yay term *rat<sup>3</sup> raap<sup>2</sup>* appears to contain *raap<sup>2</sup>* ‘to carry on the two ends of a shoulder pole.’

41. *The Nest*

Representation: 89%

Motivated: 40.0%

Thereof Analyzable: 21.0%

Thereof Colexifying: 19.0%

Thereof by Contiguity: 5.1%

Thereof by Similarity: 32.1%

Recurrent associated meanings: house/home, bird, den/lair, container/box,  
 beehive, bed, web, basket, raft, shelter, hole, grass, rubbish, egg

There is one metaphorical pattern that is more frequent than any other lexico-semantic associations found for it in the database. This is the likening of the bird's nest to the house of humans. The association can be realized by both colexification and morphologically complex terms, although the latter strategy is more common. Complex items can be most often be translated literally as 'bird house' (such as Kaluli *o:ba*: a 'bird house,' this is also the case in Efik, Kanuri, Mbum, Ngambay, Noni, Yoruba, Quileute, Miskito, and Malagasy), but some variation is found. For instance, Embera has *ĩmbaná dhe* 'bird home,' and Kiowa *tou-sq̣'n* consists of *tou* 'house' and *sq̣'n* 'grass.' The same structure is also found in Cayapa, compare also Toaripi *ori roro* 'bird rubbish/refuse/weeds' and Pawnee *rahkisiitu* /*raar-kisiit-u* / 'ITER-flat.reed-NOM.' In turn, for the association with rubbish, compare Pipil *-tapahsul*, consisting of the 'unspecified object' prefix *ta-* and *-ku:pahsul* 'rubbish' (there is an alternative possible source for the sequence *pahsul*: *mu-pahsulua* "for one's hair to be messed up"). In San Mateo del Mar Huave *omb-iüm* is analyzable as 'hole-house' (this complex term can also refer to a 'house' itself, compare colexification of 'nest' and 'hole' in Kolyma Yukaghir). Kyaka features in addition the term *anda pingi* 'house/nest root,' Kapingamarangi *hale ngogo* 'house egg,' and Sora has *asu:ɣtidən* /ə-'su:ɣ-'tid-ən/ 'POSS-hut.for.temporary.use-bird-NMLZ.' Moreover, there is a derived term in Kiliwa, and colexification with 'house' and/or 'home' is found in Hausa, Khoekhoe (where the term is formally derived from a verb meaning 'to build, construct), Kyaka (also colexifying "open valley area" and 'shed'), Muna, Badaga, Wintu, Maxakalí, and Lenakel (and in Tasmanian with 'hut, camp' more specifically).

A further metaphorical transfer is one from the meaning 'bed' rather than 'house' to the 'nest.' This is found in Samoan and Bislama by morphologically complex terms (*fa'a-moega* 'like-bed' and *bed blong pijin* 'bed of bird' respectively), and in Guaraní and Manange by colexification (similarly, Khoekhoe colexifies 'nest' with "sleeping place, resting place"). The Burarra term for 'nest' consists of the verb for 'to mound up,' *gapulawa*, nominalized by prefixation of the 'general' class marker *gun-* and also means 'clump, knot.' The same pattern is possibly also found in Wintu (however, the relevant term is only attested from one speaker). In four sampled languages, Muna, Kolyma Yukaghir, Jarawara, and Hupda, an association with 'container' or 'box' is found (in Welsh also with 'case'), in the case of Hupda by the morphologically complex term *hūtāh cá?* 'bird box,' and in Muna by a term derived from a verb meaning 'to collect, receive' inter alia. Similarly, Buin and Fijian colexify 'nest' with specific types of baskets (Fijian also with other meanings). Efik has *e'fök i'nuën* 'sheath bird.'

A very interesting association is found in Nunggubuyu and Middle-Eastern Tasmanian: in both languages, the relevant terms can also refer to a ‘raft’ (in Nunggubuyu, to a raft made from paperbark specifically), and this may be a case of provenience contiguity, since (at least some) birds in this area build their nests in paperbark trees.

Common is also colexification of ‘nest’ with other habitations of animals: with ‘den’ or ‘lair’ in fifteen sampled languages (Kwoma, Rotokas, Waris, Basque, Greek, Khalkha, Kolyma Yukaghir, Nez Perce, Wintu, Cashinahua, Maxakalí, Hani, Mandarin, Sedang, where the relevant term colexifies ‘garment, blanket,’ and Yay), with ‘hive’ in Nivkh, Hawaiian, Mandarin, and Lesser Antillean Creole French, and with ‘web’ in Buli, Muna, and Ineseño Chumash. Badaga and Hawaiian also colexify ‘shelter’ generally.

Given that the function of a bird’s nest is to lay eggs in it, it is surprising that lexico-semantic associations with ‘egg’ are quite infrequent. Alongside Kapingamarangi *hale ngogo* ‘house egg’ which was already mentioned, Ket has *engaj* /eʔŋ-àj/ ‘eggs-sack,’ and Chukchi *kətcəjɔlyen* is related to *kətcəŋjo-* glossed as ‘sit in ambush for’ in the consulted source, which also cites an older source stating its meaning to be ‘to sit on eggs’ (compare Itzaj *k’otb’aj* ‘brood, nest, cluck with chicks,’ derived from *k’ot* ‘to brood’).

Morphologically complex terms constituting semantic relations that occur only in one of the sampled languages include Swahili *kiota* ~ *kioto*, consisting of *ota* ‘sprout’ and a noun class prefix and Muna *kaofe* ~ *kaufe*, apparently derived from *ofe* ~ *ufe* to “squeeze cooked rice into a round shape, compress.” Abipón features a term derived from a verb meaning ‘to assemble,’ Fijian *vakavevêde* may contain *vaka*, which alongside grammatical functions means ‘be like, resemble,’ *veve* ‘crooked, bent’ and *dē* ‘to fix firmly.’ Derivation by suffixation of classifiers is found in two languages of the Amazon: Chayahuita *pěpětē* is derived from *pěpērīn* ‘carry’ by the instrumental classifier *-tē*, and Cubeo *cūribu* appears to be derived from *cūrō* ‘place, site’ by the classifier *-bu* for cylindrical or round objects. Khoekhoe *haires* contains *hai* ‘tree, wood, plant, stick,’ Badaga has a term for “nest among stones” containing a constituent meaning ‘stone,’ Upper Chehalis one meaning ‘moss,’ and another Upper Chehalis term contains a verb meaning ‘to settle down, reside.’ The Tuscarora term for ‘nest’ appears to contain the word for ‘mother.’ The relevant Arabela term contains a classifier for balls of fibres, the Guaraní one a constituent meaning ‘pile,’ Piro *sreta* is related by unknown means to *sure* ‘leaf, sheet of paper,’ Sáliba *juwōchē* appears to contain *juwo* ‘hairs, feathers,’ and the Great Andamanese term *ârrâm* is derived from *râm* ‘to cover’ by prefixation of a possessive marker. Semianalyzable terms including a constituent meaning ‘bird’ are attested in Khoekhoe, Kemtuiik, and Kwoma.

Unique patterns of colexification include that with ‘winnow’ as well as ‘to pour into vessel’ inter alia in Efik, with ‘iris of the eye’ and other meanings, and, by a different term, with ‘support’ and ‘husband’ in Buin, with “bush camouflage, hide” in Ngaanyatjarra, with ‘rudiment’ and other meanings in Kyaka, with “(family) goods, possessions” in Sentani, with ‘writing’ in Sko, with ‘cave’ in Abzakh Adyghe, with ‘bear’ and ‘womb’ in Bezhta, with ‘cradle’ in Kashaya, with ‘niche, nook’ in Lesser Antillean Creole French, with ‘straw’ in Wintu, with ‘base’ in Carib, with ‘diaper’ in Cavineña, with ‘hammock’ and/or ‘yawl’ in Wichí, with ‘shelter’ and ‘gathering place’ and other meanings in Hawaiian, with ‘string, rope’ and by another term with ‘blood’ in Bwe Karen, with ‘placen-

ta, afterbirth,' cocoon,' and 'handle' in Lenakel, with "body cavity, such as armpit, hollow of the knee" as well as 'tide, morning tide' in Mandarin (though the latter term has a different etymon, Pulleyblank 1991: 51), and with "shelf up high" in Sedang. There are extensions to the human sphere in Khalkha ('cell of a political party') and Mandarin ('"nest" of robbers'). Similarly, the Basque term *habia* may also be employed metaphorically with reference to the human sphere. Finally, note that Yuki *noh* ~ *no?* 'nest' and *noh* ~ *no?* 'to live' are identical segmentally.

#### 42. *The Plant*

Representation: 59%

Motivated: 66.9%

Thereof Analyzable: 20.6%

Thereof Colexifying: 46.8%

Thereof by Contiguity: 28.1%

Thereof by Similarity: 0%

Recurrent associated meanings: tree, grass/weed, thing, grow, vegetable, flower/blossom, bush/shrub, vegetation, leaf, forest, seed, to plant, sprout/shoot, green, plantation, stalk, land/earth

It is common for terms for 'plant' to be associated in some way with a more specific type of plant further down the taxonomy (autohyponymy, in terms of Horn 1984). Thus, Khoekhoe, Ngambay, Rendille, Ngaanyatjarra, Yir Yoront, Sora, San Mateo del Mar Huave, Kashaya, Kiowa (where the relevant term is furthermore identical segmentally with that for 'to be dewy'), Nuuchahnulth, Quileute, Yaqui, Ancash and Imbabura Quechua, Fijian, Hawaiian, Rotuman, Samoan, Takia, Tetun, Yay, and Bislama colexify 'plant' with 'tree' (some also with further meanings discussed in section 65), and there is an overt term exhibiting the relationship in Bora (*úmé-hé-wu* 'tree-CL.tree-DIM'). Muna, Abzakh Adyghe, Badaga, Bezhta, Ket, Kildin Saami, Upper Chehalis (by a semianalyzable term containing an element meaning 'grow'), Cheyenne, Ineseño Chumash, Embera, Guaraní, Maxakalí, and Wayampi colexify 'plant' with 'grass' (similarly, Chickasaw colexifies 'wild plant' with 'weed' and Miskito 'plant' with 'medicinal herb'). There is a semianalyzable term where the identifiable meaning is 'vegetation, weeds' in White Hmong. Meyah, Badaga, Kildin Saami, San Mateo del Mar Huave, Santiago Mexquititlan Otomí, and Central Yup'ik colexify 'plant' with 'flower, blossom' (Badaga also with 'harvest' and 'offspring'). Rendille, Kyaka, Badaga, Nuuchahnulth, Embera, and Maxakalí colexify 'plant' with 'bush, shrub.' Moreover, San Lucas Quiaviní Zapotec *cwàa'n* is ambiguous between a narrow reading 'alfalfa' and the broad reading 'plant in general.'

Muna, Yaqui, Wayampi, and Hawaiian, as well as Ngambay, also colexify 'forest' (Abzakh Adyghe also 'decoction, medicine' and, probably from there on, 'cleaning, chemical product,' Bezhta also 'hay,' and Embera also 'mountain' and 'thicket,' see also Buck 1949: 521 for similar associations in Ancient Greek and Latin).

Dongolese Nubian, Abzakh Adyghe, Ket, Kildin Saami, Central Yup'ik (dialectally), Guaraní, and Hawaiian colexify 'plant' with 'vegetable,' and Dongolese Nubian, Rendille, Cheyenne, Embera, and Guaraní with 'vegetation' generally (similarly, Hawaiian colexifies

‘greenery’ and has the analyzable term *mea ulu* ‘thing vegetation’). Embera and Kaingang colexify ‘plant’ with ‘plantation.’

Moreover, it is common that ‘plant’ is colexified with specific parts of plants (automeronymy, if one likes). Thus, Hausa, Kyaka, Carrier, Nuuchahnulth, and Tuscarora colexify ‘plant’ with ‘leaf’ (Kwoma also with ‘paper;’ in Tuscarora, the ‘leaf’-reading is archaic), and there are complex terms betraying this association in Yoruba (*ewé-ko* ‘leaf-farm’), and Nivkh (*pan'-tjomr-ku* ‘grow-leaf-PLURAL.SUFFIX’); furthermore, Carrier has the redundant complex term *yenhwoʔan*, containing *yen* ‘earth’ alongside *ʔan* ‘leaf, plant.’ Hausa colexifies ‘plant’ with ‘root’ (alongside ‘source of river’ and ‘double sheet of paper’ inter alia), and Kyaka has *renge pingi-pi* ‘stem/basis/origin root-ASSOC.’ Baruya, Upper Chehalis, Aguaruna, and Huambisa colexify ‘plant’ with ‘seed’ (Aguaruna by a term that might be semianalyzable, containing a constituent meaning ‘field’), and Ngambay colexifies it with ‘stalk, reed,’ an association mirrored by the verbal Pawnee term *tat-kus* ‘stalk-to.be.sitting.’ Nuuchahnulth *ʔaqapt* can also mean ‘branch.’

However, there are also complex terms for ‘plant’ not of the two major types so far discussed. Nivkh *pan'-tjomr-ku* ‘grow-leaf-PLURAL.SUFFIX’ was already mentioned, and there is also a number of other languages where there is an association with ‘to grow.’ Sora, for one, has *a-nʔeb-an* ~ *ne:b-an* ‘(POSS-)to.grow-N.SFX,’ Hani *ja-ssaq* ‘grow.tall-CLASS.SMALL,’ Malagasy *zavamaniry*, analyzable as /zàvatra-maniry/ ‘thing-to.grow,’ and such terms are also found in Katcha and Japanese. Abzakh Adyghe has *š°x°entʔeyacʔe* /š°x°antʔe-ye-čʔ(e)/ ‘green-ABSTR-grow,’ with the item meaning ‘to grow’ also colexifying ‘seed’ and ‘egg’ inter alia (compare the Hawaiian term mentioned above). The association with ‘growing’ is also present by a derived term in Central Yup’ik and in Ineseño Chumash by an unclear morphological process. Upper Chehalis *sʔacáytmš* contains *ʔac* ‘grow’ and *=tmš* ‘land,’ and Chayahuita has *noʔpa quëran paporin-so* ‘earth from grow-3SG.SUB’ (for this association, compare also the Carrier term mentioned above, as well as that there is a semianalyzable term with the meaning ‘soil’ as the identifiable constituent in Guaraní). Upper Chehalis has another term derived from another verb meaning ‘to plant, to grow’ colexifying ‘seed,’ while ‘plant’ and ‘it is growing’ are colexified in Acoma, and there are semianalyzable terms featuring a verb meaning ‘to grow’ in Upper Chehalis and Central Yup’ik; moreover, Kosarek Yale *bongodoba* contains *bongodob* ‘everything which grows,’ and note also the similarity between Yuki *čur* ‘plant’ and *čur-h-* ‘to grow’ (see Buck 1949: 521 for this association in Indo-European). As also noted for Indo-European by Buck, Badaga, Kolyma Yukaghir, and Tuscarora betray an association of ‘plant’ with ‘sprout, shoot’ by colexification (Badaga also colexifies ‘harvest’ and ‘offspring,’ and Kolyma Yukaghir also ‘beam’), and the association is realized by alternation of noun class in Swahili. Moreover, there are a number of terms in which ‘plant’ is associated lexically with ‘to plant.’ Yoruba has *ò-gbìn* ‘NMLZ-to.plant’ (colexifying ‘planter, farmer’), and *ohun ò-gbìn* ‘thing NMLZ-to.plant,’ Hupda *yúm* for ‘plant planted by humans’ is analyzable as ‘plant/sow.NMLZ,’ and Fijian has *i kei* ‘DERIV to.plant.’ Mandarin and Vietnamese, as isolating languages, have complex terms for ‘plant’ featuring the constituents ‘to plant’ and ‘thing,’ and there is a semianalyzable term in Piro.



As noted throughout, there are some languages in which the relevant terms feature a constituent meaning 'thing.' Rendille, Kwoma, Ngaanyatjarra, and Yir Yoront colexify 'plant' and 'thing' directly (but 'grass' is excluded from the denotational range of the Rendille term; Kwoma also colexifies 'animal'), and a semianalyzable term of this kind is also found in Efik. While Abzakh Adyghe has a complex term for 'plant' with a constituent meaning 'green,' Dongolese Nubian directly colexifies 'green, light blue' with 'plant,' and there is a semianalyzable term in Ineseño Chumash.

Other associations include: in informal usage, Khoekhoe *haii* also means 'marihuana,' and Rendille *géey* also means 'dance, song.' Kaluli *é* also denotes a 'seedling' specifically. The Basque term *landare* is derived from *landa* 'field.' Lesser Antillean Creole French *plan* also means 'plan, scheme' (due to collapse of Fr. *plante* and *plan*), Kiowa *goup* also 'vine,' and Nuuchahnulth *ḷaqapt* also 'Kinnikinnick, Bearberry.' Guaraní *yva ra'y* is analyzable as 'fruit DIM,' and *ka'avo* appears to contain *ka'a* 'mate,' which according to another consulted source alongside 'mate' also means 'plant' by itself, as well as 'vegetation.' Bwe Karen *-mu* also means 'day' inter alia, Hawaiian *mea ulu* is analyzable as 'thing vegetation,' Rotuman *ʻai* colexifies 'stiff, rigid' and other meanings, *hū* in the same language also means, inter alia, 'lower end,' and Takia *ai* also means 'pelvis.'

#### 43. *The Puddle*

Representation: 46%

Motivated: 85.6%

Thereof Analyzable: 68.1%

Thereof Colexifying: 19.6%

Thereof by Contiguity: 53.7%

Thereof by Similarity: 26.0%

Recurrent associated meanings: pond/lake, water, swamp, stand/sit/be stagnant, mud, hole, spring/well, lagoon, rain, pit

Often, terms for 'puddle' (not distinguishing between 'puddle' and 'pool') are contiguity-based complex terms of the lexical type, with one of the constituents being a word for 'water.' There is a recurrent subtype, namely that with terms meaning 'to stand, to sit' or 'stagnant' acting as the second constituent, as in Ineseño Chumash *s-qil-ilik'in* '3SG/3SG.POSS-water-stand.' Such terms are also found in Efik, Abzakh Adyghe, Pawnee, and Tetun (where 'to stagnate' is colexified with 'to form puddles'). San Mateo del Mar Huave has *ndorrop yow* 'hole water,' and a term with this structure is also found in Lesser Antillean Creole French. Sora has *rupa:luj'dan* /rupa:-l'u:ŋ-d'a-n/ 'hole-pit-water-N.SFX,' Piro *tkomha* /tkome-ha/ 'small.hole-water/eye,' and the meanings are directly colexified in Greek; similarly, Yoruba has *kòtò ketere* 'pit small.' Aymara has *uma uma* (reduplicated from *uma* 'water'), and precisely the same structure is found in Samoan. Dadibi has *pu áí ge* /pu áí ge/ 'mud water nut/egg/small.object' for 'puddle of dirty water' specifically; note also colexification of 'puddle' with 'mud' in Basque (also with 'drop,' 'waterhole,' and 'reservoir'), Khalkha, Aguaruna, Wayampi ('muddy ground' more specifically), and a semianalyzable term where the identifiable constituent is 'dirt' in Chayahuita, and one with 'earth' and another one with 'adobe' in Bora. Pipil (Cuisnahuat dialect) has *ta:l-a:pu:ni* 'ground-water-be.born,' which colexifies 'puddle' with 'swamp' and 'spring, well.' The

former pattern of colexification is attested also in Muna, Basque, Khalkha, San Mateo del Mar Huave, Arabela, and Bora, and the latter in Basque, Oneida, and Hawaiian. Mbum has *ḡi-mbàm* ‘at-rain’ and Berik *aro fo* ‘rain water/lake/river,’ while Nunggubuyu *wargalg* is glossed as “rainwater on ground (including puddles).”

Other complex terms involving a constituent meaning ‘water’ are Efik *mkpö-diök'hö-möñ* ‘thing-be.placed-water,’ glossed as “water surrounded by land, collected in a depression; a pool; a pond,” Kaluli *ho:n wakan*, where *ho:n* is ‘water’ and *wakan* the name of a “inedible wild taro-like plant that grows by water,” Ket *hýnul*, analyzable as /hýna-ül/ ‘small water,’ Cheyenne *tsé-a'kóom-oëha*, containing *tsé-* ‘that which is’ and *oëha* ‘water,’ Kashaya *ʔahqʰa šu-naca'la-w* ‘water by.pulling-remain-ABS,’ Oneida *kahnekóniʔ*, analyzable as /ka-hnek-No-ʔ/ ‘NEUT.AGENT-liquid/liqour-be.in.water/cook.in.water-STAT,’ Central Yup'ik *meqcarlluk*, perhaps containing *meq* ‘fresh water’ and *-rrluk* ‘thing that has departed from its natural state,’ Guaraní *y-no'ö* ‘water-accumulation,’ Piro *tkomha* /tkome-ha/ ‘small.hole-water/eye,’ Hawaiian *laha-laha wai* ‘RED-extended water,’ Rotuman *tān kəlu* ‘water encircle’ and *tān hāe* ‘water contain’ (these terms also mean ‘lake,’ and the latter also ‘bay, inlet’). There are semianalyzable terms in One, Chickasaw, Piro, Toba, and Tsafiki, as well as one featuring a constituent meaning ‘liquid’ in Yanomámi.

Furthermore, Efik, Baruya, Buin, Kyaka, Rotokas, Kosarek Yale, Abzakh Adyghe, Cheyenne, Haida, Lesser Antillean Creole French, Tuscarora, Copainalá Zoque, Bora, Cashinahua, Guaraní, Ancash Quechua, Hawaiian, Lenakel, Rotuman, Sedang, and Yay use a single term for both ‘lake, pond’ and ‘puddle’ (Lenakel colexifies “pool on the reef at low tide” more specifically; Lesser Antillean Creole French colexifies also ‘mast, pole’). Some of them are analyzable, their internal structure being discussed in section 34; furthermore, Yoruba has *ògòdò kekere* ‘pond small’ and there is a derived term in Hawaiian. In Copainalá Zoque and Ancash Quechua, ‘lagoon’ is colexified in addition.

Other associations include: Buin *rurugapau* also denotes a “flooding on road, gutter, stream,” Muna *tobhi* can also refer to “the deepest part of a river or the sea” *inter alia*, and the variant *šalba* of Khalkha *šalbaya ~ šalbaya ~ šalba* ‘pool, puddle, mud’ also means ‘quick, quickly.’ Bora has *adó-wa* ‘drink.NMLZ-CL.pond’ for a “well or pond in the bush where animals drink (slightly salty) water.” Embera *nāmbúa* means ‘puddle, pool’ with masculine gender and ‘profundity, depth’ with feminine gender. Yanomámi colexifies ‘puddle’ with ‘ditch,’ Great Andamanese *elākākōdo* might contain *kōdo*, ‘coil of rope’ alongside the possessive prefix *ākā-*. Hawaiian colexifies ‘puddle’ with “small pool for stocking fish spawn,” ‘cistern,’ ‘mollusc,’ and other meanings. Another term, *hāpuna*, is figuratively used with the meaning ‘child,’ and, due to English influence, ‘harpoon.’ Finally, Sedang *tóng* also is the name of a kind of grass.

#### 44. The Rain

Representation: 99%

Motivated: 28.9%

Thereof Analyzable: 8.8%

Thereof Colexifying: 21.4%

Thereof by Contiguity: 21.5%

Thereof by Similarity: 2.7%

Recurrent associated meanings: water/liquid, cloud, rainy season, storm/rainstorm, sky, rain shower, neck, day, calabash for carrying water, descend, lake/pond, weather, fall, stone

'Rain,' as one might intuitively have expected, is a meaning that is not very frequently expressed by motivated terms cross-linguistically. The most common association, by configurational contiguity, is that with 'water' (further associations due to this pattern of colexification, for instance that with 'river,' are not discussed here, though see sections 34 and 47 as well as § 6.2.2.5.). This pattern is also suggested for Indo-European in diachrony (Buck 1949: 68). Sixteen of the sampled languages, namely Hausa (which also has the optional complex term *ruwan sama* 'water/rain sky'; *ruwa* also has many other meanings, among them 'juice'), Anggor, Gurindji, Ngaanyatjarra (here also 'waterhole' is colexified), Waris, Kosarek Yale (where the relevant term also means 'life-sap, vitality' as well as 'talk, criticism'), Cheyenne, Itzaj, Pipil (colexifying also 'well' and 'pool'), Lesser Antillean Creole French (colexifying also 'sweat'), Xicotepec de Juárez Totonac, Aguaruna, Huambisa, Hupda, Jarawara, and Miskito directly colexify 'rain' with 'water.' Alternatively, a few sampled languages also feature analyzable terms that may be either characterized by contiguity or by similarity, in both cases with 'water' acting as contiguity anchor. Examples of the former are Koyraboro Senni (Labbezanga dialect) *beene-hari* 'sky-water,' Ket *ules /ül-ēs/* 'water-sky,' Bororo *bu-butu* 'water-fall/descent/birth' (compare Dadibi *tulubage*, presumably analyzable as */tulubo-ge/* 'fall.down-nut/egg/small.object,' as well as colexification of 'rain' and 'to descend' in Kiowa), and examples of the latter are San Lucas Quiaviní Zapotec *nnyi'sgyihah*, analyzable as */nnyi'ihs-gyihah/* 'water-stone' (compare also Tsafiki *suhua*, analyzable as */su-hua/* 'stone-big') and Rama *yát si* 'abscess water.' Further semianalyzable terms with 'water' are found in Bororo and Rama, and similarly, Miskito has *pura laya* 'above liquid' for 'rainwater' and another semianalyzable term featuring *laya*, while Lenakel *nihi-n* may possibly, according to the source, be analyzable as 'liquid-3SG.POSS.' Four languages in the sample, Mbum, Nunggubuyu, Yir Yoront, and Cahuilla, colexify 'rain' and 'cloud(s)' or more specifically 'raincloud' by provenience contiguity (Yir Yoront also uses this term for the 'rainbow serpent,' compare section 44), while in Kiliwa, 'rain' is *kwi y h+uhaa-k* 'cloud 3+arrive-HR,' and in Khoekhoe, the same root yields terms for 'rain, thunderstorm' and 'cloud, raincloud,' with nominal designants distinguishing the readings. In three sampled languages, Badaga, Pawnee, and Abipón, 'rain' is colexified with 'storm' or 'rainstorm' (similarly, Southeastern Tasmanian colexifies 'rain' and 'thunderstorm'), and in another three, Badaga, Arabela, and Yanomámi, relevant terms may also refer to a 'rain shower' (as is the case in Irish, Buck 1949: 68). Nivkh colexifies 'rain' with 'weather,' for which compare dialectal Central Yup'ik *cella-lluk* 'world/outdoors/weather-bad.'

In four other languages, Gurindji, Muna, Wayampi, and Hani 'rain' and 'rainy season' are colexified (in Hani also 'summer'). Bezhta and Sentani colexify 'rain' with 'day' (according to Nikolayev and Starostin 1994, this pattern is due to accidental phonological changes in Bezhta). Furthermore, two languages of South America colexifying 'rain' and 'water,' Aguaruna and Huambisa, also colexify 'calabash used to carry water' in the same

term (Aguaruna in addition, ‘juice’ and ‘mucous membrane’). Katcha *thimbido* ~ *thombodo* also means ‘sky, heaven,’ and Manange also colexifies ‘rain’ with ‘sky’ (note that according to Dixon 1982: 69, in dialects of Dyirbal a cognate means ‘sky’ in northern dialects and ‘rain’ in related languages, so this pattern may be more frequent cross-linguistically). Itzaj and Jarawara colexify ‘rain’ with ‘lake.’ Kapingamarangi and Samoan, presumably by homonymy, colexify ‘rain’ with ‘neck.’

Other associations include: the Buli term *ngmoruk* also rarely refers to ‘thunder, lightning,’ while *wen-zuk*, perhaps analyzable as ‘sky-head’ means ‘up, upwards, above’ normally, but is used for ‘rain’ by the section of the clan that is responsible for rain and which hence must not utter the ordinary term. For Koyraboro Senni *baana* compare *baana*, meaning ‘to be soft?’ Rendille colexifies ‘rain’ with ‘God,’ while Burarra *yorr* is also the name of a type of shellfish and Meyah *mós* also means ‘fish’ generally. Kosarek Yale *mok*, dialectally meaning ‘rain,’ also without dialect restrictions means “place, area” and “side sprout, offshoot.” Abzakh Adyghe *šx(e)* can also mean ‘to plaster, to fuse,’ Chukchi *il(ə)il* is connected (reduplication?) to *ilə* ‘damp,’ Greek *vrochí* is connected to *vréchō* ‘to dampen,’ Japanese *ame* with different prosodic properties also denotes ‘candy, sweets,’ and Nivkh *lyx* also means ‘rainy’ and ‘wet.’ Nuuchahnulth *kiic* ‘light rain’ also denotes “raining mist, spray,” while Yana *bariku* ~ *bareeku* contains *ba-* ‘to spill, to flow’ and *-ri(ku)* “down, downhill, on the ground.” Bororo *butaodoge* is also the name for spirits causing rain. Embera *kúe* means ‘rain’ with feminine gender and ‘heavy downpour’ with neuter gender. The denotational range of the Jarawara term *isi/iso* includes ‘leg, lower leg,’ ‘handle,’ ‘stalk,’ ‘rain,’ ‘hasta,’ and ‘seedless fruit,’ while Macaguán *em* also means ‘winter.’ As a verb, Piro *hina* also means ‘to come,’ Sáliba *oxo* also means ‘leaves,’ and Wichí *iwumcho* inter alia contains *wu* ‘make’ and the locative suffix *-cho* ‘under.’ Bwe Karen colexifies ‘bug’ and other things, while Lesser Antillean Creole French *lapli* is also used figuratively with the meaning ‘shower.’ Malagasy *òrana* also denotes the ‘crayfish,’ and Mandarin *yu3* also means ‘and.’

#### 45. The Rainbow

Representation: 90%

Motivated: 33.7%

Thereof Analyzable: 28.8%

Thereof Colexifying: 6.3%

Thereof by Contiguity: 3.4%

Thereof by Similarity: 30.8%

Recurrent associated meanings: snake, bow/arc/bend, sky, god, rain, snare, color, thunder, rope, cloud

The ‘rainbow’ is a concept predominantly expressed by metaphor-driven terms. At the same time, it is also a meaning for which these conceptualizations very frequently exhibit what seem to be areal patterns, both on a large and on a small scale. The two most common associations are on the one hand that with terms meaning ‘arc’ or ‘bow’ or terms meaning ‘bend’ or ‘bent,’ which is most common in the Old World, and on the other hand that with ‘snake,’ which is most common in the Americas and New Guinea. The former pattern is attested in Greek (*ouránio tóxo* ‘heavenly/relating.to.sky bow/arc’), Kildin Saami

(*tīr'mes'-jūkks* 'thunder-bow'), Welsh (*bwa'r Drindod*, as well as *bwa'r arch*, both containing *bwa* 'bow' and the latter term *arch* 'ark'), Kiliwa (*s-?+nwaaw=x-u?+siw=h+qhaa-tay* 'INST.LONG-DN+bend=CAUS-OBL+??=3+shoot-FREQ'), Yuki (*si'k* was 'blue/green bent'; the analysis is considered questionable by lexicographers, there also is the alternative term *sikwástlik* /*si'k-wástlik*/ 'blue/green-stand'), Kaingang (*ta vyj* 'rain bow,' alongside *ta no* 'rain arrow'), and Tetun (*arku iha lalehan* 'arc have sky' and *arku-iris* 'arc-sky,' both apparently calqued from Portuguese). Moreover, Rendille colexifies 'rainbow' with "arc(s) of stones" for rituals and "barrier of stones," Swahili *upinde* consists of *pinda* 'bend' and a noun class prefix, and there are semianalyzable terms in Badaga and Quileute (where the relevant term probably contains an element meaning 'bent'). Interesting to note is also that Buli *nagortom*, a loanword from Twi *nyankopon-ton*, is folk-etymologized to *Naawen gori tom* "god has made a bow." In fact, associations between 'rainbow' and 'thunder,' as betrayed in Kildin Saami, are an identifiable areal pattern of Eurasia themselves, also occurring in the sample in Ket *aqqot* /*ekŋ-qo't*/ 'thunder-path' (compare also Cashinahua *navan bai tapia*, containing *nava* 'dance, singing,' *bai* 'path/river,' and *tapi* 'firefly') and Nivkh *lyj petr* 'thunder ornament' (*petr*, more specifically, is the name for a multicoloured ornament worn on shoes); the term also means 'ulcer,' 'sore,' or 'wound' (by virtue of them changing colors when healing?). The phenomenon is discussed in Räsänen (1947), see also § 6.4.3.5.

The other major association is that with 'snake' (or a specific snake species), occurring by colexification in Burarra (Gun-nartpa dialect), Nunggubuyu, Toaripi (where *lavai* is at the same time the name of a particular snake species as well as 'tortoise' and 'dolphin' inter alia), Yaqui, Bororo, Jarawara (also colexifying 'jungle monster'), and by analyzable terms in One (*suwol ilwola* 'snake shadow'), Kosarek Yale (*mano yame* 'snake soul/image,' this term is said to also denote the 'spirit of the snake'), Kashaya (*muša'laqol*, analyzable as /*muša'la-ʔahqol*/ 'snake-tall'), Aguaruna (*págki wajáu* 'boa resting'), and Rama (*shírking núngkit* 'boa throat'); semianalyzable terms exist in Kaluli, Kyaka (where 'snake' is colexified with 'grub' and other like creatures), and Chayahuita. For Kwoma, the source notes that "[t]he rainbow is often identified as saliva spat out by a snake ..." There is a well-known mythological complex in cultures of Australia revolving around the Rainbow Serpent (see contributions in Buchler and Maddocks 1978), and this is evidenced by colexification of 'rainbow' with 'snake' in Nunggubuyu and with specific snake species in Burarra (and note colexification of 'rain,' 'raincloud,' and the 'rainbow snake' in Yir Yoront; there are semianalyzable terms where the identifiable constituent is 'rain' in Khoekhoe and Bislama). Mead (1933) shows that beliefs of the Rainbow Serpent also occur in New Guinea, and points to several similarities between the myths of the Arapesh and those typical of Australia, and Brumbaugh (1987) reports on the Rainbow Serpent as represented in the mythology of the Feranmin and other Mountain Ok groups, also noting that it "in behavior and attributes ... corresponds closely" (Brumbaugh 1987: 32) to the ethnographic evidence from Australia. Neither of these authors makes a very strong case for continuity between the beliefs in New Guinea and Australia, although this position appears to be implicit at least in the title and discussion of Mead (1933). A remarkable parallel is Toba, where *quemoxonalo* ~ *qamoxonalo* 'rainbow' (containing *nquemoxon* 'grasps violently') also denotes a mythological great viper which punishes by causing an earth-

quake if a menstruating woman looks for water (the Marsalai, the Rainbow Snake in Arapesh mythology, pursues a menstruating woman in a story reported by Mead 1933: 41, and there are taboos forbidding menstruating women to frequent places associated with Marsalai according to Mead 1933: 43, and in Feranmin mythology the movements of the Magalim in the earth causes earthquakes, according to Brumbaugh 1987: 27).

A pattern that is more clearly due to areal spread is the association between ‘rainbow’ and ‘snare’ in North America. In the sample, this is found by colexification in Cheyenne (here, the relevant term also conveys the meanings ‘fishhook,’ ‘fishing line,’ and ‘fishing pole’) and Lakshota, and by the complex term *hoḵwez-pil* ‘cold-snare’ in Carrier. For Lakshota, the consulted source remarks: “The Indians believed that the rainbow caused the end of a rainstorm by trapping it, so that no more rain could get through” and Hall (1997: 56) states that “ ‘snare’ or ‘trap’ was a common Plains name for the rainbow because rain disappeared when the rainbow appeared.” Note also that in Tuscarora, the same root *-wene-* yields terms for both ‘rainbow’ and ‘iron.’

There are, of course, also lexico-semantic associations which do not betray any clear areality in their distribution. These include the association with ‘color’ in Kiowa (*tsoue-kuat* ‘water-painted,’ literal translation by lexicographer: “many-colored”) and Yanomámi (*ōmayari no mayō* ‘evil.spirit color footprint’). The association is formally realized by colexification in Wintu, where *sa-q* means ‘colored,’ ‘colors,’ and ‘rainbow’ and is in turn probably related to *sa-q* ‘to bleed’ (Wintu in addition features complex term on the basis of *sa-q*, and note also the redundant Mandarin term *cai3-hong2* ‘color-rainbow’), that with ‘rope’ in Baruya (*byaangwila*, literally ‘light rope’) and Blackfoot (*náápiwa otó’piim ~ náápiwa otokáa’tsis* ‘Naapi’s rope,’ Naapi being the name of the trickster and creator god), that with ‘cloud’ in Chukchi (*celgia-jaik* ‘red-cloud’) and Hawaiian (*ao akua* ‘cloud god,’ compare also the Rendille children’s term *irtiyyó=hí Waahk* ‘beads=POSS god,’ Chickasaw *Chihooowa inaahpisa* ‘which contains *chihooowa* ‘god’ and *nannahpisa* ‘promise,’ and Malagasy *antsiben’andriamànitra /àntsi-bè-n-andriamànitra/* ‘knife-big-GEN-god’), with ‘moon’ also in Chickasaw (*ninak ontoomi* ‘moon shine’), and at least by a semianalyzable term in Yoruba (Chickasaw, in addition, also has other semianalyzable terms presumably containing *hashi* ‘sun, moon;’ one of them colexifies ‘rings around the moon’). In addition, conceptualizations in which ‘rain’ and ‘sky’ act as contiguity anchors are frequent. For ‘rain,’ these include Khoekhoe *tū-lhana-b* ‘rain-garden-3SG.MASC’ (which is restricted to Bible translations) and *lgao-lhana-b* ‘stop.raining-garden-3SG.MASC,’ Mbum *mbàm-pélé* ‘rain-tomorrow,’ Abipón *oah-eta* ‘rain-AGT,’ and the Kaingang terms already mentioned above (furthermore, semianalyzable terms in Anggor and Kwoma contain an identifiable constituent meaning ‘rain’). For ‘sky,’ they include Basque *ortzadar ~ ostadar /ortzi-adar/* ‘sky-horn,’ Laz *m3a-ort’apu* ‘sky-belt’ (maybe the association with ‘belt’ is an areal pattern of the Caucasus, given that Bezhta *mašola* is perhaps borrowed from Georgian *ašuni* ‘belt’), Kolyma Yukaghir *kužu:n-šöril’ə* and *kužu:d-oḡora* ‘sky-tongue’ (as well as *jukud-onora* ‘small tongue’), the Haida “story word” *qwii sdal* ‘sky slope,’ Tuscarora *yurehyáhuks*, containing the roots *-reḡhy-* ‘sky’ and *-huk-* ‘to light up’ and colexifying ‘Aurora Borealis’ and ‘Milky Way,’ Wayampi *iwa-ləwa* ‘sky-on,’ Yanomámi *hetu kēki shīl*, consisting of *hetu* ‘sky,’ *shīl* ‘light’ and the collective quantal classifier *kēki* (see § 4.4.1.1.), Fijian *drō-drō-lagi* ‘run.away-

RED-sky,' as well as the associations with 'bow' in Greek, where 'sky' acts as a contiguity anchor. Moreover, there is a semianalyzable term where the identifiable constituent is 'sky' in Lesser Antillean Creole French.

As becomes clear from the above list, the conceptualization in each language is widely different, in spite of the common semantic element acting as the contiguity anchor. Unsurprisingly, there are even more unique and culture-specific metaphorical conceptualizations for this concept. These include: Hausa *bakan gizo*, which also denotes a "single arch in a roof," contains *baka* 'mouth, bow' and *gizo* 'mythical spider,' while Dongolese Nubian *káǵibbē* is literally 'killer of horses.' Buin *iroro(na)* is also an "epithet for males in songs." The Kwoma term *wariipoy* "green of rainbow, 'liver' of rainbow" is also used for a 'type of small tree that grows near trees in forests.' The Rotokas term *govugovuto* seems to contain *govugovu* 'clean out, purge.' Khalkha *solungya* also denotes the 'weasel' or the 'Siberian marten.' Sora *ilij'bo:ŋən ~ 'ilim'bo:ŋən* is built around the noun root *'bo:ŋ-* 'class of deities,' but the additional material remains obscure. Cahuilla *píyaxat* also denotes a 'worm with two horns.' Comanche is unique in having a term for the 'rainbow,' *pisi ma?rokóo?*, that is literally 'infected thumb.' The Itzaj term, colexifying "fog rising from earth" is *kis witz* 'fart hill.' For Lake Miwok *káccakaca* 'rainbow,' compare perhaps *kacáakaca* 'bluebird, *Sialia mexicana*,' a multicolored bird with blue, gray, and orange plumage. Pawnee *huraahkipic* is semianalyzable: it contains *huraar* 'be land.' Yana *laki-yaa* is analyzable as 'navel person,' the term means 'newborn baby' in Central Yana and 'rainbow' in Northern Yana. The Central Yup'ik term *agluryak* is derived from *agluq* 'ridgepole, center beam of a structure' by means of addition of the postbase (cf. § 4.4.2.) *-yak* 'thing similar to,' and San Lucas Quiaviní Zapotec *garre'ed bihih* is literally 'cart air/wind,' a variant of the term is *garre'ed gyeht*, probably analyzable as 'cart squash/pumpkin.' Bora *tuúhi* might contain the classifier for small rivers *-hi*, Bororo *jure* colexifies 'rainbow' with 'dance' as well as 'queue,' the Cubeo term *náme* is also used to refer to a 'string of liana,' while Ancash Quechua *turmanyay* also denotes 'gas that emanates from the earth.' Fijian *mudu* or *valemudu* denotes a half rainbow (*mudu* means 'cut off, ceased, ended,' *vale* is 'house'), Hawaiian *ānuenuē* is also used to refer to "the scallop-like design on tapa and tapa-beater," and *haka 'ula a kāne*, a poetic term for the 'rainbow,' is analyzable as 'perch red poss Kāne.' White Hmong *zaj-sawv* is literally 'dragon-rise,' and Tetun *baur* colexifies 'rainbow' with 'to cheat, swindle,' and *namerak* with "to grow murky." Sedang *kia pōtea* apparently contains *kia* 'ghost,' Vietnamese *cầu vồng* is analyzable as 'bridge curved,' while Yay *roj<sup>5</sup> ?wa<sup>1</sup>* contains *roj<sup>5</sup>* 'bright.' The San Mateo del Mar Huave term *ndequiamb poj* seems to contain *poj* 'terrestrial turtle.' Finally, Piro colexifies 'rainbow' with 'pus' and the meaning to be treated in the following section: the 'resin.'

#### 46. The Resin

Representation: 66%

Motivated: 49.5%

Thereof Analyzable: 27.0%

Thereof by Contiguity: 5.0%

Thereof Colexifying: 24.6%

Thereof by Similarity: 34.2%

Recurrent associated meanings: water/liquid/juice, tree, blood, tar, stick/sticky, glue, milk, honey, wax, syrup, rust, molass, semen, cud, phlegm, birdlime, pus

Most frequent for this meaning (for which ‘sap’ and ‘gum’ were accepted as proxies if it was clear that indeed sap of plants and trees is meant) are lexical associations with ‘water’ and/or more generally ‘liquid’ or, less generally, ‘juice.’ One or more of these meanings are colexified in Lavukaleve, Bezhta, Ket, Sora, Maxakali, and Hawaiian (Bezhta also colexifies ‘drink’ and Ket also ‘alcoholic beverage’). More common, however, are terms in which one of these meanings merely is one of the constituents in analyzable terms, as in Nivkh *tiyr-čox* ‘tree-juice.’ Alongside Nivkh, such terms with ‘tree’ as the other constituent are also found in Kanuri, Japanese, Cheyenne, Chickasaw, Lesser Antillean Creole French, Yaqui, Carib (colexifying ‘torch’), Imbabura Quechua, Mandarin, White Hmong, and the alternative where ‘bark’ rather than ‘tree’ is used as the contiguity anchor is attested in Piro (*mta-ha* ‘bark water’). A semianalyzable term in which one constituent can be identified to mean ‘water’ is found in Kyaka, and further variations of this pattern are Bororo *bato kuru* ‘mangaveira liquid’ and Ancash Quechua *hacha-pa wiqi-n* ‘plant-GEN juice/tear-3SG.’ In Santiago Mexquititlan Otomí, the relevant term is *dehe* ‘yo-mu(n)hño ngizá /dehe ‘yo-mu(n)hño ngi-zaa/ ‘water walk-good sap-tree.’ The language also has another term for ‘resin,’ namely ‘ba zaa ‘milk tree.’ Associations with ‘milk’ are also found by colexification in Kwoma, Sora, and Ineseño Chumash.

An association by colexification between ‘resin’ and ‘blood’ is attested in Kwoma, Cubeo, Maxakalí, Fijian, and Hawaiian. Also note Sko *hí* ‘sap’ and *hì* ‘blood’ as well as that Jarawara colexifies ‘blood’ with ‘red sap of certain trees’ specifically and, by another term, also colexifies ‘puss,’ which is probably an error in the source for ‘pus.’ At any rate, ‘pus’ and ‘sap, resin’ are colexified in Sora and Piro (alongside ‘rainbow’ in the latter language). Bislama has *blad blong tri* ‘blood of tree.’ Colexification with ‘honey’ is found in Nez Perce, Oneida, and Bwe Karen (which also colexifies ‘to be pure, clean’ inter alia). Three sampled languages, Tuscarora, Hawaiian, and Tetun, colexify ‘resin’ with ‘wax’ (similarly, Rotuman colexifies ‘sealing wax’ specifically), Nez Perce, Oneida and Tuscarora colexify ‘syrup,’ Nez Perce and Tuscarora also ‘molasses,’ and Nez Perce also ‘sorghum.’ Two of these languages, Tuscarora and Hawaiian, also use the relevant term for ‘tar,’ an association also occurring in Efik, Aguaruna, and Bora; the Tuscarora term *uθrèweh* also colexifies ‘cement’ as well as ‘jam’ and ‘jelly,’ and the Hawaiian one also ‘printers’ type’ and ‘sinker on a fishing line’ inter alia. Nez Perce furthermore colexifies ‘cud chewing’ (by another term than that participating in the above patterns), and ‘resin’ and ‘cud’ are also colexified in Chukchi. Four other languages, Buli, Waris (where the relevant term *nénel* appears to contain *né* ‘forest, forest product’), Fijian, and Tetun colexify ‘resin’ with ‘glue,’ either by provenience contiguity if resin is actually used as glue, or by perceptual similarity based on their common stickiness. In fact, there are languages in which the words for ‘resin’ explicitly make reference to this. For instance, Hawaiian *pīlali* denotes the “[h]ardened sap ... of the kukui tree, gum; resin, birdlime; wax” but also means “gummy, sticky” inter alia, and the association with ‘sticky’ is also found in Berik and Wayampi by colexification. Welsh has *defnydd gludiog o coed* ‘matter sticky of tree,’ Lenakel *nouanehapwiit*, containing *noua* ‘fruit’ and



*apwiit* ‘stick, cleave to,’ and a semianalyzable term in which the meaning ‘sticky’ or ‘stick’ figures occurs in Kapingamarangi. Furthermore, there are a number of mostly metaphor-driven complex terms for ‘resin’ in which ‘tree’ acts as contiguity anchor. These include Ngambay *nən-kake* ‘tear-tree,’ San Mateo del Mar Huave *aonts xiül* ‘excrete tree,’ Guaraní *yvyra ry’ái* ‘tree sweat,’ Piro (*gagmuna-*)*kshi* ‘(tree-)rainbow,’ Hawaiian *hū lā’au* ‘swell tree,’ Malagasy *tsironkàzo*, analyzable as /*tsìro-hàzo*/ ‘taste-tree/wood,’ Sedang *chhá lóang* ‘split tree/twig,’ as well as Takia *ai pat-an* ‘tree kidney-3sg.’ There are semianalyzable terms with the identifiable constituent meaning ‘tree’ in Kosarek Yale, Cavineña, and Sáliba. Sahu and Hawaiian colexify ‘resin’ with ‘birdlime.’ Two languages of Mesoamerica, Itzaj and San Lucas Quiaviní Zapotec, colexify ‘resin’ with ‘rust’ (San Lucas Quiaviní Zapotec also with ‘nectar’ and ‘lymph’). Toba and Hawaiian colexify ‘resin’ with ‘phlegm,’ and Miskito and Hawaiian with ‘semen.’

Other associations include: Gurindji *tinung* also denotes the “bloodwood sap” specifically, Kosarek Yale *keles* ‘resin used for glueing and tightening a drumskin,’ according to the consulted source, might be related to *kel* ‘female’ and *es* ‘flower,’ and *widi*, also a name for resin used for tightening a drumskin, also denotes a variety of sugar cane. Abipón *liciRa* is derived from *-ici ~ -icir-* ‘assemble, connect,’ while Arabela *riya-ca* appears to be analyzable as ‘star:PL-CLASS,FRUIT.’ Bororo colexifies ‘resin’ with ‘rubber,’ and Cayapa also means ‘crescent in a river.’ Chayahuita *yaqui* is derived from the verb *yaquirin* ‘to cut well or completely’ by means of suffixation of the classifier for liquids, *-i*. Huambisa colexifies ‘resin’ with ‘tattoo,’ and Kaingang *jênjo* might contain *jên* ‘to eat.’ Guaraní *aysy* also means ‘gluten.’ Rama *shubli ~ ubli ~ yubli* can also refer to a ‘stain’ or a ‘secret.’ Hani *ziq* also means ‘hemp,’ a relevant Hawaiian term colexifies ‘to roll, turn’ inter alia, and Rotuman *pulu*, which can refer to “any adhesive substance” in general, can also be used with the meaning ‘seal.’ Hawaiian *pīlali* may contain *lali*, meaning ‘greasy’ inter alia. Finally, probably accidentally, Tetun colexifies ‘resin’ with ‘candlenut’ and ‘to prune, clip,’ and Yay *θa³* also denotes “any of various devices having a wheel.”

#### 47. *The River*

Representation: 96%

Motivated: 43.6%

Thereof Analyzable: 11.0%

Thereof Colexifying: 33.2%

Thereof by Contiguity: 29.2%

Thereof by Similarity: 2.6%

Recurrent associated meanings: water, spring/well, lake/pond, juice, valley, river bed, ocean, way/trail/track, big, rain, flow/run, channel/drain/ditch/gully, floodwater, Milky Way

The most frequent association in this case, clearly, is with ‘water.’ Many languages in the sample colexify ‘river’ (or ‘stream,’ ‘creek,’ etc., which were accepted as proxies) with ‘water,’ and many of them also use the general term for ‘water’ also for other bodies of water, such as a ‘lake’ or a ‘spring.’ These will not be discussed here, see sections 34, 44, and 56, and especially § 6.2.2.5. for discussion. Terms which colexify ‘water’ (or ‘fresh water’ specifically) and ‘river’ are found in Efik (also colexifying ‘tide’), Mbum, Ngambay,

Anggor, Baruya, Berik, Buin (where the relevant term is also a “general name for all ruta designs” as well as an epithet for a “plumb child”), Kyaka, Mali, Muna, Meyah, One, Sko (where also other meanings are colexified), Waris, Kosarek Yale (where the relevant term also means ‘life-sap, vitality’ and ‘talk, criticism’), Abzakh Adyghe, Badaga (also colexifying ‘floodwater,’ as is the case in Ngaanyatjarra and Abzakh Adyghe), Cahuilla, Upper Chehalis, Haida, Kashaya, Pipil, Quileute, Wintu, Copainalá Zoque, Bororo, Cayapa, Chayahuita, Guaraní, Huambisa, Jarawara, Kaingang, Macaguán, Tsafiki, Wayampi, Yanomámi, Bislama, White Hmong, Sedang (where further apparently unrelated meanings are colexified), and Takia (40 languages all in all and thus a little more than twenty-five per cent of the sampled languages). Kosarek Yale, Abzakh Adyghe, Kashaya, Copainalá Zoque, Tsafiki, and Bislama colexify also ‘juice,’ and sometimes still further meanings.

Some of the languages just mentioned also have special dedicated terms for ‘river’ alongside the colexifying term, such as Efik, which has *akpa* ‘river’ alongside the general term *mōñ*. Also, some languages have optional analyzable terms, such as Abzakh Adyghe, where *psə* colexifies ‘water,’ ‘juice,’ and ‘river,’ for which latter *psə-x°e* may also be used (*x°e* is glossed as “être, devenir, advenir, augmenter, mûrir” in the source), or Wintu, where *mem* means both ‘river’ and ‘water’ and *bohem mem čuha* is used for river, *bohe* meaning ‘to be big’ and *čuha* ‘to flow.’

Alternatively, there are also complex terms for ‘river’ where one of the constituents is ‘water,’ and also here there are cross-linguistically recurrent patterns, and in fact, the optional Wintu complex term already points to two of them: complex terms where the second constituent is ‘big’ are found in Noni, Santiago Mexquititlan Otomí, Yana (this term also refers to the Sacramento River in particular), and Yaqui, where *batwe* is perhaps analyzable as /ba’a-bwe’u/ ‘water-big.’ Malagasy *renirāno*, analyzable as /rèny-rāno/ ‘mother-water’ is very likely an instance of the pattern as well, given the augmentative function of terms for ‘mother’ in many languages (Matisoff 1992) and the corresponding grammaticalization path; compare also Wappo *?éču tùča* /*?éču túča*/ ‘creek big.’ In some Indo-European languages, terms for river go back to a term for ‘water’ or more specifically, ‘flowing water’ (Buck 1949: 41). Central Yup’ik *kuik* is said to contain the “deep root” *ku* ‘flowing liquid,’ and terms for ‘river’ in which the notions of ‘flowing’ or ‘running’ figure in addition to ‘water’ are found in Fijian, Samoan, and Tetun (where ‘stream’ and ‘torrent’ are colexified; note also the redundant term *y-syry* ‘river/water-flow’ in Guaraní). In one sampled language, the word for ‘river’ and ‘flowing water’ is of the derived type, namely Cahuilla *wáni-š*, derived from *-wáne* ‘to flow.’

‘Way,’ ‘trail,’ or ‘track’ as the second element in complex terms is found in Laz, Carrier (where ‘track’ is colexified with ‘trace, vestige’ and ‘site’), and Piro (note also Wayampi *ia-la-pe* ‘canoe-of-way,’ and that Lengua *thlinga wathuk* ‘stream’ contains *thlinga* ‘movement,’ as well as that Kashaya *bí?da* contains the root *?da* meaning ‘extend, stretch’ which is also found in the term for ‘road,’ *hi?da*). This association is realized formally by colexification in Toaripi (“because the rivers and creeks are the highways through much of the low lying Elema countryside;” the language also colexifies ‘manner’ and ‘method,’ compare section 92, and, uniquely, ‘hand’ and ‘arm’), and Cashinahua. Moreover, Lesser Antillean Creole French *lawivie* could be analyzed as containing *lawi* ‘street,’ but it is more

likely to go back to French *la rivière*. Itzaj has *ok ja* 'leg/foot water/lake/rain' (compare the Yay term *ka<sup>1</sup> ta<sup>5</sup>* 'leg river,' denoting a "very large river" specifically), Pawnee *kic-ka* 'be.liquid-on.horizontal.surface' (this term also denotes the Arkansas river specifically; alongside Pawnee, there are also other languages in which the term for 'river' is ambiguous between 'river' in general and a particular river: Dongolese Nubian *úru* also refers to the Nile specifically and assumes the meaning 'to wash out, rinse' as a verb, Badaga *gangi* ~ *gange* also to the 'Ganges' and the goddess Parvati, Bororo *oroaribo* also to the Rio Paraguay and the Rio São Lourenço in Mato Grosso alongside its capability to refer to a certain spirit, and Mandarin *he2* also to the Huanghe), Maxakalí *kônāgkox*, analyzable as /kōnā'āg-kox/ 'water-hole,' Hawaiian *kaha-wai* 'water-place' and *muli-wai* 'after/last-water' (which also means 'estuary'), and Manange *māfan 2kju* 'low.river.valley water' (the association with 'valley' is also found in Khalkha, Nez Perce, Bora, Huambisa, and Hawaiian by colexification. Note further the possible etymological connection between Basque *ibai*, from earlier *\*hibaie*, with *ibar* 'valley'). Similarly, Kwoma, Badaga, Itzaj, and Rotuman colexify 'river' with either 'channel,' 'drain,' 'ditch,' or 'gully.' Badaga also colexifies 'bottomland, lowest spot, depression,' and Mandarin also 'plain.'

Also note the similarity between Tasmanian *liapota* 'river' and *liena* 'water,' leading Plomley (1976: 372) to connect the two. Furthermore, Bwe Karen has *chí-bú* 'water-in,' and semianalyzable terms in which one constituent is 'water' exist in Piro, Rama, and Kapingamarangi.

There are also associations exclusively realized by colexification in the languages of the sample: Buli, Hausa (alongside some highly specialized culture-related meanings), Khalkha, San Lucas Quiaviní Zapotec, and Hawaiian (by the analyzable term *kahena wai* 'flowing water') use the same term for 'river' and 'river bed,' and in five languages, Efik, Ngambay, Kashaya, Arabela, and Bora, 'river' or 'big river' and 'ocean' (but with the exception of Efik and Kashaya not 'water'!) are colexified. Waris, Kosarek Yale, Pipil, and Jarawara colexify 'water' and 'river' with 'rain,' while Yir Yoront and Ancash Quechua colexify 'river' with 'Milky Way' (for complex terms for the Milky Way involving a constituent meaning 'river' compare section 37). In Ket, 'water,' 'liquid,' as well as 'alcoholic beverage' are colexified, and in White Hmong, 'water,' 'river,' and 'wine' are.

Other associations include: Efik *uquā* can also refer to a 'flood,' and *akpa* in the same language also means 'first.' Hausa *kogi* is also used as an epithet as well as the name of a children's game, and Khoekhoe (Haillom dialect) *dommi* also means 'throat, voice.' Ngambay colexifies 'waves.' The Burarra term *angartcha* is derived from the verb *gartcha* 'be stuck' by prefixing of the noun class marker *an-*. Muna *laa* also means 'stem, stalk' and 'straight' inter alia, and another Muna term, *oe*, also is used metaphorically for "interest (in money)." Meyah *mei* also means 'sperm' (alongside 'water'). Rotokas *gae* 'waterway, river' has a verbal reading "follow a course, heed talk, drift, wander." Badaga *ole* ~ *hole* can also mean 'reservoir' and 'swamp,' while Basque *ibai* can also refer metaphorically to an 'enormous lot.' The Khalkha term *yulduril* is derived from the verb *yulduri-* meaning 'to spill, to be poured out, to pass through' (alongside other meanings colexified) by means of the abstract nominalizer *-il*. Chickasaw *abookoshi* contains *oshi* 'son,' a morpheme widely used in this language to form metaphorical expression usually conveying a meaning of

smallness. Kiowa colexifies ‘river’ with ‘moon’ (the relevant term is also a name for a game). The term for ‘river’ in Santiago Mexquititlan Otomí, *hñe*, is also the word for ‘mirror’ (as well as for ‘to put on’). Wintu *mem*, colexifying ‘river’ with ‘water,’ also means ‘wet,’ ‘thirst,’ and ‘to baptize.’ Aguaruna *namák(a)* also means ‘fish.’ The Macaguán term *pemnát* also means ‘tube,’ and Rama *ri* is also used adjectivally with the meaning ‘wet.’ Hani *lolbaq* contains *lol*, a classifier for rivers, and *baq* means ‘direction’ or ‘thin’ and acts itself as a classifier for the side of a mountain and pages of books. Mandarin *he2* also means ‘peace, harmony’ (with different etyma, Pulleyblank 1991: 122) *inter alia*, *xī1* ‘small river in mountains’ also ‘to suck in,’ ‘knee’ and ‘tin’ (all reflecting different etyma, Pulleyblank 1991: 328–330), and *chuan1* also “to bore through, pierce” (both identical segmentally already in Early Middle Chinese, Pulleyblank 1991: 60). Sedang colexifies ‘large river’ with ‘to imprison,’ and Yay with “to put (ones’s own money, goods) with another’s larger amount” and ‘to listen, hear.’

#### 48. *The River Bed*

Representation: 26%

Motivated: 61.5%

Thereof Analyzable: 40.2%

Thereof Colexifying: 25.2%

Thereof by Contiguity: 35.9%

Thereof by Similarity: 7.7%

Recurrent associated meanings: river/stream, valley, water, way, flow, base/basis, deep, hole/hollow, place

Lexico-semantic associations for this concept are manifold. Frequently, either ‘river’ or ‘water’ is one of the constituents in analyzable terms, though note that five sampled languages, Buli, Hausa, Khalkha, San Lucas Quiaviní Zapotec, and Hawaiian (by the analyzable term *kahena wai* ‘flowing water’) directly colexify ‘river, stream’ with ‘river bed’ (Nez Perce furthermore with ‘waterway’) and five others, Sentani, Abzakh Adyghe, Khalkha, Lengua, Pawnee (by a term literally translatable as “enclosure on a surface”), and Rotuman, colexify it with ‘valley’ or ‘ravine’ (Abzakh Adyghe also colexifies ‘precipice,’ and Rotuman also ‘gutter, gully’ and ‘channel, trench’), while the Khalkha term has still other meanings and *inter alia* rarely also assumes the meaning ‘large lake.’ In Sentani, the relevant term *jaba* ‘dry river bed’ appears to contain *ja* ‘sink,’ for which compare Toba *ca’amgue* ‘dry riverbed’ with *ca’amgui* ‘sunk, to lower in the middle.’ Otherwise, like for ‘river’ itself, there are associations with ‘way,’ ‘road,’ or ‘track’ realized by analyzable terms in Mali (*arenggi atha iska* ‘river her road’), Nez Perce (*wé’le?skit* /*wé’le-ʔiskit*/ ‘flow way;’ for this term compare also Wintu *me’m ɕ’oh-i* ‘water/river flow-NOMINAL.STEM.FORMANT’), and Miskito (*tingni bila* ‘river way’ and *li bila* ‘water way;’ *bila* colexifies many more meanings alongside ‘way’). Also note in this context Guaraní *ysry-ha*, probably analyzable as /*ysry-(a)ha*/ ‘river go’ as well as Nivkh *er myy dif* ‘river go.downstream track’ which is only perhaps so analyzable, but would fit this pattern. Two languages of Eurasia, Ket and Kolyma Yukaghir, have terms for ‘river bed’ containing an element with the meaning ‘deep’ (*sést hóbaj*, containing *sēs* ‘river,’ *hòq* ‘deep,’ and *ba’ŋ* ‘place,’ and *unuy-čeginmā* ~ *unuy-čiginma* ‘river depth’ respectively), while in two Austronesian languages, Fijian and Tetun, associa-

tions with 'hole' or 'hollow' are encountered (*dreke-ni-wai* 'hollow/cavity-POSS-water' and *mota-kuak* 'river-hole' respectively). In addition, an association with the meaning 'base' is found by colexification in Hawaiian (alongside other colexified meanings, the relevant term *papakū* probably contains *papa* 'flat surface, reef, layer' and *kū* 'to stand,' both constituents also have other meanings) and by the analyzable term *isalẹ odò* 'bottom/base river' in Yoruba. The metaphorical transfer from 'bed' to 'river bed' appears to be peculiar to Europe, occurring in the sample only by the analyzable term *gwely afon* 'bed river' in Welsh. There are also languages in which body-part metaphors are employed to convey the meaning 'river bed': 'guts' and 'river bed' (as well as 'inside') are colexified in Jarawara, Khoekhoe has the analyzable term *lā-lnā-b* 'river-stomach/interior-3SG.MASC,' and Rendille colexifies 'arm, hand,' 'elephant's trunk,' and 'river bed.' Tetun *mota-fatin* is analyzable as 'river-place,' compare the semianalyzable Carrier term *nethayikēt* containing *kēt* 'place,' that the second element in the Cavineña term *ejiri quini* means 'broad place,' and that Pawnee *huukahaaru*?, denoting a 'dry riverbed' specifically and at the same time colexifying 'beach, shore' and "dry bed in a pond," is literally "in water place." Pawnee also has the term *huukaahaaru*?, containing *huuka(wi)*- 'along a stream course' and *haar* 'place,' this term also means 'beach, shore.' Other morphologically complex terms with 'river' include Efik *isōñ akpa* 'earth/ground river,' Rotokas *uuko gae*, containing *gae* 'river' and *uuko*, 'liquid state, fluid,' 'to collect water,' Sora *əluŋ'nai* /ə-luŋ-'nad/ 'POSS-sleep.or.soak.in.water-river,' Chickasaw *abookoshi' shila-* 'river dry-NMLZ,' denoting a dry river bed specifically. Hawaiian colexifies 'riverbed' with 'bottom' among other meanings. Furthermore, Samoan *alitivai* contains *vai* 'water' but is not amenable to full morphological analysis, the same situation is encountered in Piro. There is a semianalyzable term featuring a constituent 'river' in Yuki. Finally, Sedang *chúa* is also the name of a weaving design inter alia.

#### 49. The Root

Representation: 97%

Motivated: 32.6%

Thereof Analyzable: 6.9%

Thereof Colexifying: 25.8%

Thereof by Contiguity: 5.4%

Thereof by Similarity: 8.4%

Recurrent associated meanings: base/basis, tendon, origin, trunk,

reason/cause, vein, nerve, muscle, stump, tuber, foot, leg, branch, buttocks, *lia-na*, root of tooth, tongue root, stalk, foot of hill, tendril, thread, tree

The 'root' is a meaning expressed in many languages by non-motivated, non-analyzable terms. If they are motivated, however, body-part metaphors are used in some languages to express the concept. In three sampled languages, Buli, Burarra, and Miskito, 'root' is colexified with 'foot,' and in Abzakh Adyghe, *lapse* is analyzable as /λ(e)-ps(e)/ 'foot-string.' In Buli, Ngaanyatjarra and Yir Yoront, 'root' is colexified with 'leg' (in Buli, by the same term that also colexifies 'foot,' it also means 'branch' inter alia, as does the relevant Ngaanyatjarra term which denotes a 'side root' specifically. Buck 1949: 522 also reports this association for Indo-European. Yir Yoront also colexifies 'tail' of certain fish species

and is used generically for certain molluscs). Bwe Karen has *θo-kha-wi* ‘tree-leg-vein.’ There are also some languages in the sample in which ‘root’ and ‘vein’ are colexified, namely Rendille, One, Basque, Carib, Lengua, and Miskito (similarly, Gurindji more specifically colexifies ‘single root of tree’ with ‘vein,’ and Hawaiian ‘small root, rootlet’). However, for this association, cross-linguistic tendencies in the directionality of the mapping are hard to assess, given that in Kanuri, ‘vein’ is *zâr bû-bè* ‘root blood-of.’ Many sampled languages also colexify ‘root’ with ‘tendon.’ This association is found in Kwoma (where also ‘tendrill’ is colexified, a pattern shared with Kiliwa), One, Chickasaw, Ineseño Chumash (also meaning, presumably by provenience contiguity, ‘bowstring’), Pawnee, Jarawara, Lengua, Miskito, and Hawaiian. In addition, ‘root’ and ‘muscle’ are colexified in Koyraboro Senni, Ineseño Chumash, Jarawara, and Hawaiian (where the relevant term conflating all three meanings is also used figuratively for ‘womb’ and ‘offspring’), and ‘nerve’ in Koyraboro Senni, Basque, Ineseño Chumash, Miskito, and Hawaiian. Note that there is a large overlap between the four groups due to the fact that very frequently ‘tendon’ and ‘vein,’ and somewhat less frequently also ‘muscle’ and ‘nerve’ are colexified (see sections 141 and 147). The Koyraboro Senni and Highland Chontal terms also denote a ‘liana, vine’ or species of liana. Two languages, Sora and Rotuman, colexify ‘root’ with ‘root of tooth,’ Hausa and Yoruba colexify it with ‘buttocks,’ and another two, Cashinahua and Mandarin, with ‘tongue root.’ Sko *hanging* might be based on a metaphorical comparison with *háng* ‘end of intestines, kidney’ (the second element of the term is unknown, and a similar situation is encountered in Kwoma; for this possible association, compare Great Andamanese *ârçhâg* ‘root’ with *ôngchâg* ‘kidney,’ presumably motivated by alternation of noun class for which compare discussion in § 4.1.1.2.). Lavukaleve *kala* also means ‘collarbone,’ and for Sentani *kambu*, compare *kambi* ‘neck.’ The Wintu term *çaraw ɣosi* contains *çaraw* ‘green land, field, valley,’ and *ɣosi* might be related to *ɣOs* ‘fog, steam, gas, lungs.’ Kaingang *jã-re* appears to be analyzable as ‘tooth-field,’ and Lenakel *nuk-* may also refer to the ‘armpit.’

Intra-domain ties are somewhat less frequent. Seven languages, Buli, Abzakh Adyghe (by the analyzable term mentioned above, note that this term also means ‘thread’ and compare Swahili *mzizi*, analyzable as /m-uzi/ ‘3/4-thread’), Yaqui, Lesser Antillean Creole French, Abipón, Miskito, and Tehuelche colexify ‘root’ with ‘trunk’ or parts thereof (Miskito also with ‘pole’ and ‘substance,’ and Tehuelche also with ‘handle of knife,’ ‘claw,’ and ‘behind of’). Yaqui does so also with ‘stump,’ an association it shares with Biloxi, Lesser Antillean Creole French (where the term also means ‘stem’) and Nez Perce. The Hausa, Badaga, Pipil (Santo Domingo de Guzmán dialect), and Jarawara terms are also used with the meaning ‘tuber.’ Badaga *ga:su* is also used specifically with reference to the ‘potato,’ and Tuscarora *uhné?reh* is extended to ‘turnip’ and ‘vegetable’ in general. Another association is that between ‘root’ and ‘stalk,’ found in Abzakh Adyghe and Badaga (in the latter language, by a semianalyzable term containing a root meaning ‘to grow,’ ‘fertile,’ and ‘yield’ inter alia, and itself colexifying “water dripping from roots”).

There are also a number of associations where the terms for ‘root’ are used to map some more abstract meaning. Buli, Hausa, Kyaka, Abzakh Adyghe, Khalkha, Kolyma Yukaghir, Ancash Quechua, Fijian, and Malagasy colexify ‘root’ with ‘base, basis’ (Kolyma Yukaghir also with ‘custom’ and ‘similarity’), ‘origin’ is a secondary meaning of ‘root’-

terms in Buli (where it is restricted), Angkor (where the term *ahasahari* is semianalyzable containing *ahari* ‘stem’ and colexifies ‘clan’), Basque (where one of the relevant terms also means ‘root in mathematics,’ ‘stock, lineage,’ ‘tentacle,’ ‘udder,’ ‘hinge,’ ‘inclination, tendency,’ ‘beam, sunbeam’ and ‘segment’), Khalkha (also expressing the notions of principalilty and originality, among others), Nez Perce, Lesser Antillean Creole French, Embera, Rotuman, and Tetun. ‘Reason’ or ‘cause’ is colexified in Buli, Nez Perce, Hawaiian, Malagasy, Mandarin, and Rotuman (the relevant Buli term participating in these patterns is the same that also colexifies ‘trunk,’ but a different one from that realizing the associations with ‘leg’ and ‘root’ reported above, the Hawaiian term also means ‘foundation’ inter alia, and the Rotuman term is also used with the meanings ‘principal village, capital,’ ‘lower end,’ and others). Another metaphorical association, that with ‘foot of hill/mountain,’ is attested in Abzakh Adyghe and Greek.

Further less systematic associations include: Buli *kiri* can also refer to the ‘character’ of a person, ‘type,’ as well as ‘under, below’ (compare the colexification of ‘root’ with ‘bottom’ in Hawaiian, and that with “upward from below” in Kapingamarangi). Efik *āduñ* appears to be derived from *duñ* ‘to dwell, inhabit.’ Hausa *saiwa* can also refer to a “drabbish-coloured goat,” and another Hausa term, *tushe*, colexifies ‘root of plant’ with ‘plant itself’ inter alia. For Khoekhoe, it is noted that *!nomab* also means ‘root’ in the linguistic sense. Oddly, Ngambay *ndirà-kake* seems analyzable as ‘hard.to.eat-tree,’ while *ko*, another term for ‘root’ in the same language, also means ‘to cry’ and ‘seed,’ and *ngirà* can also be used adjectivally with the meaning ‘hard.’ Noni *gbwey* also denotes the ‘stinger of a bee,’ and Rendille *híy* also means ‘sour milk’ and ‘relatives, kin.’ Baruya colexifies “fine root in or out of ground” with “fat from a pig’s stomach,” while Berik *tiskar* and Dadibi *ni pedali* ‘tree root’ apparently contain *ti* and respectively *ni* ‘tree, wood.’ Kwoma colexifies ‘root’ inter alia with ‘wrinkled,’ and Gurindji *wirnturru* also means ‘horn.’ The Muna term *paraka* is also the name of an evil spirit that eats people. Meyah *ofóm* also means ‘ripe.’ The Rotokas term *vavu-rupa-to* is analyzable as ‘bitter-dark-SG.M,’ while Sahu *utu’u* also means ‘buttress’ specifically. A Badaga term for ‘root,’ *be:ru*, contains the verb *be:* ‘to grow’ and colexifies ‘finger.’ Khalkha *yndysy(n)* also means ‘race, nationality’ and is the name for the religious writings in the ‘Tantra,’ and Welsh *gwreiddyn* is also used with the meaning ‘stock.’ Chickasaw *ishtaahikki’ya* is analyzable as /isht aa-hikki’ya-/ ‘with LOC-stand-NMLZ’ (the term refers to above-ground roots specifically), Itzaj colexifies “hateful, angry,” Wintu ‘herb charm’ by a semianalyzable term containing a constituent referring to level land and *če’k* colexifies ‘ropelike root(s)’ with ‘rope, cord string,’ ‘to tie a rope.’ The Yuki term *koot’kin* ~ *kutkin* might be derived from the verb *koot* ‘to start.’ A Central Yup’ik word for ‘root,’ from the Norton Sound dialect, *acilquq*, is derived from the noun *aci* ‘area below, area under’ by suffixation of the postbase (see § 4.4.2) *-quq*, meaning ‘one that is like’ (there is the variant *acipluk* /aci-lluk/ ‘area.below/area.under-bad’). Another Central Yup’ik term, *nemernaq*, consists of the noun *nemeq* ‘binding, wrapping’ and the postbase *-naq* ‘one like,’ this term also denotes the ‘lamprey.’ Cashinahua *tapun* might contain *tapu* ‘platform, shelf, table,’ and Cayapa *telele* ‘root’ *te* ‘firewood.’ Chayahuita *itë* appears to be analyzable as /i’-të/ ‘water-INSTR,’ while Macaguán *petakomét* appears to be analyzable as /peták-omét/ ‘container-sun/moon.’ Piro colexifies ‘pelvic bones,’ and Maxakalí

*mĩmyĩpxatit* contains *mĩm* ‘wood.’ Toba *lpa’a’* ~ *lpa’a’q* is apparently related to *pa’a’*, which is the name of a plant the root of which causes hallucinations. Fijian *vũ* also means ‘cough, to cough’ and ‘to wash, cleanse,’ and Hani *alqil* and *daoqqil* might be related to *qil* ‘be firm, be durable.’ Figuratively, Hawaiian *a’a* also means ‘womb, offspring,’ as well as “to send greetings of love; joyous hospitality; joy at greeting a loved one.” Finally, Samoan *a’a* also means ‘connection, “involve, implicate,” and ‘influence,’ Sedang *rei* also is the name of a bird inter alia, while Yay *raak’* also means ‘to pull, drag.’

#### 50. *The Seed*

Representation: 89%

Motivated: 43.3%

Thereof Analyzable: 6.4%

Thereof Colexifying: 36.9%

Thereof by Contiguity: 5.7%

Thereof by Similarity: 13.1%

Recurrent associated meanings: fruit, egg, offspring/descendants, eye, testicle, semen, plant/sow, round/round thing, nut, bead, germ, heart, stone, breed, bullet, bone, plant, grow, kind, bud, face, product, peer/friend, tuber, berry, flower

Associations between ‘seed’ (or ‘grain’ / ‘pit’) and other small roundish objects abound cross-linguistically. Languages of the New Guinea area (but also some others) are particularly fond of colexifying ‘seed’ with other products of plants or animals that have the property of being round. Buin, Kyaka, Rotokas, Kosarek Yale, Abzakh Adyghe (inter alia), Wintu, Hawaiian, and Takia colexify ‘seed’ with ‘egg,’ and Gurindji, Kwoma, Kyaka, Ket (colexifying also ‘pine nuts’ specifically), Kapingamarangi, and Takia colexify ‘seed’ with ‘nut.’ More generally, ‘seed’ is associated with ‘fruit’ in many languages as well, namely in Efik (here also with ‘flower’ by a term derived from a verb meaning ‘to peel’ inter alia; in Waris, ‘flower of trees’ is specifically colexified), Ngambay, Dongolese Nubian, Buin, Kwoma (here also with “edible part of a plant”), Kosarek Yale, Khalkha, Kiowa (colexifying also ‘vegetable’ and ‘bread’), Abipón, Miskito, Piro (where the relevant term *xi* also acts as a diminutive marker), Rama (colexifying ‘peanut’ specifically), Sáliba (also colexifying ‘bud,’ as is the case in Efik and Lesser Antillean Creole French), Wayampi, Hawaiian, Takia, Sedang (colexifying also ‘pellet in blow gun’ and ‘muscle’), and Yay. Similarly, Tetun colexifies ‘berry,’ while Cheyenne *hestáhame* is literally ‘heart berry.’ Takia, alongside all these meanings, also colexifies ‘shell.’ This association may be based on meronymy in some cases, but perceptual similarity might also be at work, since many fruits are globular as well. In Rama, the term colexifying ‘seed’ with ‘fruit’ is *kat up* ‘tree eye,’ and this illustrates that associations with other round objects of small size also transcend domain boundaries. Pawnee is another language with a complex term for ‘seed’ (in particular ‘planting seed’) on the basis of ‘eye’ (*rak-kirik-u’* ‘wood-eye-NOM’), and the association is realized by colexification in Anggor, Burarra, Cahuilla (also colexifying ‘face’), Chayahuita (also colexifying ‘design of a waistband’), Jarawara (here also colexifying ‘face,’ ‘pile,’ ‘color,’ ‘end,’ and ‘pellet’), and Lengua, while the Nunivak Island dialect of Central Yup’ik and Huambisa have seminanalyzable terms. For Burarra *mipila*, the dictionary gloss reads as follows: “eye, or anything suggestive of an eye by virtue of being small and round or exud-



ing fluid, hence hail, seed, bullet, hook of woomera, glass tumbler, nipple, spring of water,” suggesting that ‘eye’ is the dominant reading (see also § 6.2.3.1.). The extension to ‘bullet’ is also found in Ngambay, Toba, and Yanomámi (here also ‘almond’ is colexified); note that ‘gunshot’ is colexified in Lengua. Four languages, Muna, Badaga, Welsh, and Maxakalí have terms colexifying ‘seed’ with ‘(small round) stone,’ and four languages in the sample, Sentani, Kosarek Yale, Nez Perce, and Samoan colexify ‘seed’ with ‘heart’ (Nez Perce also with ‘pith’). In addition, Cheyenne *hestáhame* literally means ‘heart berry,’ which suggests a metaphorical transfer of the position of the heart in the middle of the chest to the position of the seed in the center of a fruit, corroborated by the frequent association between ‘heart’ and ‘middle’ (see section 117, though note that the transfer seems to go in the other direction in Hupda *hã’wíg* which probably goes back to *\*haŋ-wíg* ‘breath-seed’). Transfer from ‘seed’ to the domain of body-parts also occurs with another meaning: seven sampled languages, Welsh, San Mateo del Mar Huave, San Lucas Quiavini Zapotec, Guaraní, Toba, Hawaiian and Kapingamarangi, colexify ‘seed’ with ‘testicle’ (Yanomámi with ‘penis’), and about the same number of languages have complex terms for ‘testicle’ based on ‘seed’ (compare section 142 as well as Brown and Witkowski 1981 for further evidence). Colexification with ‘bead’ is found in six sampled languages, Basque, Lesser Antillean Creole French, Central Yup’ik (which also colexifies ‘single fish egg’ and “any other seed-like thing”), Imbabura Quechua, Maxakalí, and Vietnamese. Moreover, there are also some sampled languages where a term for ‘seed’ in fact is glossed as having a secondary meaning of ‘round thing’ directly. These are Berik (the term seems to be semianalyzable, containing an element meaning ‘tree, wood’), Muna (where the relevant term curiously acts as a classifier also for “any big object;” this term means also ‘to be present, gather together’), One, Hupda, and Hawaiian (note also that Badaga *gunđu* assumes the meaning ‘round’ and ‘stout, strong’ when used as an adjective). Less frequent patterns involving colexification of ‘seed’ with a roundish object include that with ‘marble’ in Buli, with ‘nipple’ in Gurindji, with ‘tuber’ (as well as ‘electric battery’ and other meanings) in Kyaka and Hawaiian, with ‘flower in bud’ in Sáliba, with ‘bulb’ in Hawaiian, with ‘clitoris’ in Muna (cf. the extension of ‘fruit’ to ‘clitoris’ in Australian languages, Austin et al. 1976), with ‘tablet’ in Ngaanyatjarra, and with ‘moon, month’ *inter alia* in Rotuman.

Other recurrent associations are that with ‘bone’ in Yir Yoront, Basque (here, also with ‘fishbone’), Khalkha, and Wintu (common features: hardness and the fact that both ‘seed’ and ‘bone’ are found within larger structures, the body and the fruit respectively? at least the former is suggested by the fact that ‘hard, solid’ alongside ‘unyielding’ is also colexified in Wintu), that with ‘kind’ in three languages of different families of Africa, Efik (here only as of animals and vegetables), Hausa, and Koyraboro Senni (note that Hawaiian *‘ano’ano* might be reduplicated from *‘ano*, one of the meanings of which is ‘kind, type’), that with ‘offspring, descendants’ (sometimes also ‘child’) in Hausa, Khoekhoe, Dongolese Nubian, Kosarek Yale, Khalkha, Wintu, Hawaiian (in both Khalkha and Hawaiian also with ‘product;’ see Buck 1949: 505 for the same extension in Swedish), and Bwe Karen, which colexifies also ‘nine.’ Similarly, Badaga *bittu* colexifies ‘father’s line, patrilineage,’ as well as ‘yield’ and ‘crop’ with ‘seed,’ and the Guaraní term colexifying ‘semen’ and ‘testicles,’

*ta'yi*, is analyzable as /ta'y-i/ 'son/clot-DIM.' Still other recurrent patterns of colexification include that with 'breed' (e.g. of animals) in Efik, Basque, Miskito, and Tsafiki (similarly, Khalkha inter alia colexifies 'race, family, clan'), that with 'peer, friend' in Koyraboro Senni and Kyaka, and that with 'semen' in Dongolese Nubian, Basque, Greek, Welsh, Wintu, Aymara, and Guaraní (note that both referents are similar in that they are part of the reproductive system, and see also Buck 1949: 505 for Indo-European, in particular Swedish). Two of these languages, Greek and Guaraní, as well as Kyaka, Lesser Antillean Creole French and Hawaiian, also colexify 'germ;' in the context of the association with 'semen,' note also Kyaka *wai*, glossed as 'germs, spark of life' alongside 'seed' (as an adjective, *wai* means 'introduced, not native, not local'). Ineseño Chumash '*ami*'n also means 'body' and 'flesh, meat' alongside 'seed,' and Hawaiian colexifies "meat as in 'opihi shell or 'alamih crab."

Entirely different structures are found in terms for 'seed' when they are primarily conceived of from the point of view of agriculture. Then, terms derived from or containing verbs meaning 'to plant' or 'to sow' are common (as is the case diachronically in Indo-European, Buck 1949: 505). Koyraboro Senni and Badaga colexify these meanings, and fully analyzable terms with such structure are found in Ngambay, Abzakh Adyghe (colexifying 'cereal' generally), Cheyenne, Nez Perce, Xicotepec de Juárez Totonac, and Aguaruna, which latter for instance has *ajákma-mu* /*ajákmat-mu*/ 'sow-NMLZ.' Semianalyzable terms suggesting such a structure are also attested in Koyraboro Senni and Dongolese Nubian, and the association is recoverable diachronically in Greek. In connection with this pattern, note also Carrier *hananelyih-i* 'grow.again-REL' and Nivkh *vandu oxt* 'grow powder,' and colexification of 'seed' and 'to grow' in Basque and Abzakh Adyghe inter alia. A similar account is available for Quileute *katsàkólwa*, containing *kátsa*- 'to bury;' moreover, Baruya, Aguaruna and Huambisa colexify 'plant' and 'seed,' and Upper Chehalis colexifies 'seed' with 'plant' as well as with 'garden.'

Other associations are: Buli *biri* also means 'counter' as of a particular game, Hausa *iri* also 'slips,' and Ngambay *kànde* also 'to produce' as well as 'genetic inheritance' inter alia. Ko, another term from the same language also means 'to cry' and 'root.' Koyraboro Senni *dumari* also means 'sprout,' and in Rendille 'seed' and 'tear, teardrop' are colexified. Baruya *wia* also means 'seedling,' and Kwoma *siik*, alongside 'seed,' 'nut,' and 'fruit,' may also refer to a 'clot of blood' inter alia. Kyaka *kapa* not only participates in the colexification patterns with 'nut,' 'egg,' and 'peer,' but also may refer to the 'core' or 'nucleus' of something in general, as well as to a 'larva,' and 'fat, suet, grease' inter alia. Kosarek Yale *wana* can also refer to a 'flower-stalk.' The Muna term *lumu* also means "seedpod inside fruits" and 'moss,' and One *tala* also 'tree stump' and 'grasshopper,' while Abzakh Adyghe colexifies 'seed' with 'fur, feather' and 'tooth,' and the Khalkha term *koryngge(n)* also is used inter alia to refer to 'yeast' and 'property, resources.' Biloxi *su* also means 'blown out, extinguished,' while Upper Chehalis *smámis* is derived from the verb root *náma*- 'done, finished, quit.' Kiowa *dq̣ugā't* is literally translated "inside one," Oneida colexifies 'seed' with 'oats,' while Santiago Mexquititlan Otomí *nda* also denotes 'cotton' specifically. Pawnee *riikactikiisu'* is analyzable as /riikac-rikiis-u'/ 'crookneck.squash-kernel-NOM' (the term is also used generically for 'seeds' of all kinds), and Wintu colexifies

'seeds' with 'teeth' by one term and with 'stab, pierce, poke' by another. Carib *epipo* is derived from *epi(li)* 'flower' and means also 'stem' and 'stick,' Cayapa *ñi* also means 'flame, fire' (probably due to accidental homonymy), Kaingang colexifies 'seed' with 'braid, bolt,' Miskito with "bulk, mass, lump, particle" inter alia, Ancash Quechua with 'black and white mottled,' Tehuelche with 'leaf,' and Wayampi with 'foot' inter alia and by another term with 'almond.' Finally, Yanomámi *mo* colexifies 'seed' and 'penis.'

### 51. *The Shadow*

Representation: 92%

Motivated: 49.3%

Thereof Analyzable: 7.7%

Thereof Colexifying: 42.4%

Thereof by Contiguity: 7.2%

Thereof by Similarity: 37.8%

Recurrent associated meanings: soul/spirit/ghost, reflection/image/mirror, image/picture/drawing, photograph, shelter, dark/darkness, cold/cool, cloud

Very frequently, 'shadow' (or 'shade') is associated with 'soul,' 'spirit' or 'ghost' lexically (or, in this case, more likely primarily culturally or mythologically). This association is particularly common in Oceania and the Americas (though reported by Buck 1949: 62 to be common in Indo-European "from Homer on"), occurring by colexification in Mbum, Efik, Ngambay, Angkor (where the relevant term also denotes a type of fly), Burarra, Kaluli, Kwoma, Kyaka, Lavukaleve, Rotokas, Toaripi, Sentani, Sko, Kosarek Yale, Greek, Kolyma Yukaghir, Blackfoot, Cheyenne, Pawnee, Quileute, Wintu (colexifying also 'glimpse' inter alia), Abipón, Aguaruna, Arabela, Carib, Cayapa, Chayahuita, Jarawara, Kaingang, Maxakalí, Miskito, Ancash Quechua, Rama, Wayampi, and Lenakel (35 languages), and by complex terms in Meyah (*efená órka* 'spirit carry'), the Nunivak Island dialect of Central Yup'ik (*tarenraq* /*tarneq-aq*/ 'soul/spirit-thing.that.resembles.in.some.respect'), and Fijian (*yalo-yalo-na* 'reflection-POSS,' with *yalo-yalo* reduplicated from *yalo* 'soul, spirit'). Tsafiki, in addition, has a semianalyzable term containing *o'cô* 'evil spirit' and colexifying 'firefly.'

There are a number of further associations which cluster strikingly in Oceania and more specifically in New Guinea. Colexification with 'mirror,' '(mirror) image,' or 'reflection' is attested in Hausa (*inuwa*, the relevant term, is also a "name given to any one called Muhammadu"), Baruya, Burarra, Kaluli, Kwoma, Kyaka, Lavukaleve, Mali, Muna (where the relevant term is also the name of a tree species), Nunggubuyu, Rotokas, Sentani, Toaripi, Kosarek Yale, Itzaj (colexifying also 'nature'), Lake Miwok, Central Yup'ik (Nunivak Island dialect), Arabela, Bora, Bororo, Cayapa, Kaingang, Piro, Rama, Yanomámi, Bislama, Hawaiian (also colexifying 'embryo,' 'newly hatched fish,' 'knuckles' and 'joint,' inter alia by one of the relevant terms, and 'bright,' 'dazzling,' 'white' and similar meanings by another), Lenakel, Mandarin, Rotuman, Samoan, Sedang, and Yay, in which latter the relevant term also means 'to shine, sheen' (33 languages); moreover, in Fijian, the association is by the analyzable term mentioned above. Colexification with 'picture, image' and/or 'drawing' is attested in Bakuéri, Efik (by the term *ñwet* ~ *ñwed* derived from *wet* meaning 'paint, write, mark' inter alia, the derived term also meaning 'pattern,' 'inscription' as well as 'writing' and, presumably by further semantic extension, 'book' inter alia),

Ngambay, Noni, Rendille, Burarra, Kwoma, Mali, Nunggubuyu, One, Rotokas, Toaripi, Sahu, Sko (by the term, *bàleng*, presumably analyzable as /bà-lèng/ ‘person-hide.self’), Kosarek Yale, Yir Yoront, Abzakh Adyghe, Central Yup’ik (Nunivak Island dialect), Abipón, Bora, Maxakalí, Miskito (also colexifying ‘resemblance’), Sáliba (also colexifying ‘appearance’ and ‘color’), Fijian, Hawaiian, Lenakel, Rotuman, and Samoan (28 languages), and with ‘photograph’ in Ngambay, Noni, Rendille, Kwoma, Mali, Rotokas, Sko, Khalkha (colexifying also ‘age’ and ‘apoplexy’), Central Yup’ik, Arabela, Bora, Jarawara, Lenakel, Rotuman, and Samoan (15 languages, note that with the exception of Khalkha and Jarawara, languages which betray this association actually are a subset of languages which colexify ‘picture’).

In contrast, six sampled languages reveal a different conception of ‘shadow’ in that the relevant terms bear an association with ‘dark’ or ‘darkness’ (as also evidenced by an etymological connection between German and Greek in Indo-European, see Buck 1949: 63). This is found in Mali, Ngaanyatjarra, Kashaya (which also colexifies “sickness caused by fear”) by colexification, while realized by complex terms in Ket (*qon-sal* ‘dark-night’ and *qonij ba’ŋ*, containing *qon* ‘dark’ and *ba’ŋ* ‘place’) and the Santo Domingo de Guzmán dialect of Pipil (*ku:yuwa* ‘tree-dark’); moreover, a not entirely analyzable term containing *k’ə̀* ‘be dark’ is found in Kiowa. There is also colexification of ‘shadow’ with ‘(to) shelter,’ as encountered by colexification in Buli, Ngaanyatjarra (where the relevant term also denotes a ‘metal canopy’ as well as rings around the moon which are said to indicate coming rain), Welsh, Kapingamarangi (where the relevant term also means ‘behind some protective cover’) and Rotuman, as well as by the analyzable term *mo’ã-ha* ‘cover/protect-AGT’ in Guaraní (see Buck 1949: 62 for diachronic evidence from Irish; the Guaraní term colexifies also ‘defense’ and ‘protection’). What the structure of these terms show is that here terms are unlikely, as opposed to the examples above, to refer to the shadow of a person specifically. In fact, for instance Efik distinguishes the two lexically: *ukpõñ* is used for the shadow of things that move, and *mfut* for the shadow of things that do not. A similar distinction is made in San Mateo del Mar Huave, although terms for both variants share the same root, with that for persons or animals being inalienably possessed. Baruya has unrelated lexical items for the shadow of clouds and the shadow cast by other entities.

Two other languages, Buli and Arabela, have associations between ‘shadow’ and ‘cold’ or ‘cool place’ (note also Embera *kūrásare* ‘shadow’ and *kūrása* ‘cold;’ this association is recoverable etymologically in Donolghese Nubian, see also Buck 1949: 63 for evidence from Baltic and Serbo-Croatian). Moreover, Bororo colexifies ‘shadow’ with ‘cloud,’ and Biloxi *si natci* appears to be analyzable as /si natci/ ‘feet cloud’ (compare also Chickasaw *hoshontikachi* ‘shadow’ with *hoshonti* ‘cloud,’ and see Plomley 1976: 383 for the possibility of this connection in Tasmanian).

Other unsystematic associations include: the Buli term *yogsum*, colexifying ‘shelter’ and ‘coolness,’ also means “danger, fear, fright, dread, apprehension, terror.” Efik *mfut* is apparently derived from *fut*, meaning ‘to swell, boil, foam’ inter alia, Katcha *bogo* also means ‘place,’ and Koyraboro Senni *bii* (related diachronically to *bibi* ‘black’) is also used to refer to an ‘umbrella’ and also means ‘yesterday.’ Kaluli colexifies ‘shadow’ with “reverberation in forest or in memory,” Kwoma *mayi* also means ‘map’ and is furthermore used to refer to a variety of supernatural powers, One *iloula* ~ *ilwola* also denotes a “meas-

uring stick,' Sahu *gu'dumini* also means 'life strength,' and Western Tasmanian perhaps colexifies 'shadow' with 'fog.' Badaga colexifies 'shadow' with 'shape,' and Basque *itzal* also means 'prison' and 'respect, prestige' inter alia. Kolyma Yukaghir *numet* also means 'fontanel.' Santiago Mexquititlan Otomí *xudi* also means 'morning,' while Oneida colexifies 'shadow' with 'movie, show' (similarly, Fijian colexifies 'film'). Tuscarora *uti?θrēhsteh* is analyzable as /u-ti?θ(e)r-ē-(a)hst-eh/ 'NOUN.PREFIX-overhang-fall-NMLZ-NOUN.SUFFIX,' Bororo *arodu* also means 'bad, false,' 'robbery,' 'sign of disease and death,' as well as 'hole,' Cayapa *aama* also means 'weapon' (presumably due to collapse of Span. *alma* 'soul' and *arma* 'weapon'), and Cashinahua *baka* also means 'scar' as well as 'spouse' and 'friend.' Chayahuita *sanohuan* is derived from *sano* 'quiet' by means of a classifier suffix meaning 'one who has,' while Guaraní *kuarahy'ã* contains *kuarahy* 'sun.' Ancash Quechua *qitqi ~ qetqi* 'shadow of cloud' also means 'soot,' Yanomámi *noreshi* is also used to refer to the "phosphorescent shining of certain plants when decomposing" as well as a "double of humans incarnated in animals." Great Andamanese *ôtlêre* is apparently derived from *lêre* 'black beeswax.' Samoan *ata* also means "copy, duplicate," and, in the plural, "[l]ight and shade effect," while Tetun *mahon* also means "influence, sway" or 'framework.' Yay *ram*<sup>5</sup> is also the name of a 'big hawk that soars in the sky and eats chicken,' Vietnamese *bóng* also means 'ball,' 'cotton,' and 'flower,' whereas Bislama *sado* (< Engl. *shadow*) also means "humiliation, dishonour."

## 52. The Sky

Representation: 97%

Motivated: 49.1%

Thereof Analyzable: 11.8%

Thereof Colexifying: 37.9%

Thereof by Contiguity: 17.1%

Thereof by Similarity: 22.7%

Recurrent associated meanings: heaven, high/above/up, cloud, blue, top, air, god, weather/climate, day, light, world, rock, sun, roof, rain, ceiling, hole/opening

The most frequent lexico-semantic association for 'sky' (or 'firmament') is that with 'heaven(s)' by colexification, occurring in 44 languages, namely Buli, Hausa, Khoekhoe, Ngambay, Swahili, Yoruba, Buin (colexifying also 'swollen' inter alia), Dadibi, Kaluli, Kyaka, Ngaanyatjarra, Rotokas, Sentani, Toaripi, Kosarek Yale, Yir Yoront, Badaga, Basque, Chukchi, Ket, Khalkha, Kildin Saami, Highland Chontal, Kiliwa, Lesser Antillean Creole French (which also colexifies 'paradise'), Lake Miwok, Lakota, Pawnee, Pipil (Santo Domingo de Guzmán dialect), Quileute, Central Yup'ik, San Lucas Quiaviní Zapotec, Carib, Maxakalí, Piro, Hani, Hawaiian, Lenakel, Malagasy, Rotuman, Samoan, Tetun, Yay, and Vietnamese (similarly, Muna colexifies 'sky' with "space between heaven and earth").

The second most frequent association, that with the meanings 'high,' 'above,' or 'up, upward,' has a very striking pan-American distribution in the sample, while being rare elsewhere in the world. This is unexpected, given the seemingly related grammaticalization path 'sky' > 'up' reported by Heine and Kuteva (2002: 279) also attested in languages of other regions of the world. In the sample, the association is realized by colexifi-

cation in Buli, Mbum, Rendille, Upper Chehalis, Ineseño Chumash, Nuuchahnulth, Jarawara, Lengua, Macaguán, and Miskito, but occurs also frequently by morphologically complex terms: Kashaya has *qali q<sup>h</sup>aʔbe* as ‘clear/above rock,’ Lake Miwok *līle-wali* ‘high-world,’ Santiago Mexquititlan Otomí *ma-hets’i* ‘loc-high,’ Pipil (Santo Domingo de Guzmán dialect) *ka-ihakhu* ‘in/at-high,’ Aymara *alaji-pacha* ‘above-whole,’ Piro *ten-hohne* ‘high/tall/deep-expansive,’ and Imbabura Quechua *jawa pacha* ‘above space.’ Moreover, Comanche has the formally redundant term *tomobaʔatʰ* /*tomo-paʔatʰ*/ ‘cloud/sky-above,’ and a semianalyzable terms suggesting such a structure is encountered in Ineseño Chumash.

Several of the associations by terms of the lexical type reported above also occur in other configurations elsewhere. Associations with ‘cloud’ (which are also frequent in Indo-European, Buck 1949: 52-53) occur by colexification in Yoruba, Lavukaleve, possibly in Nunggubuyu, Bezhta, Upper Chehalis, Cheyenne (with doubts on behalf of lexicographers), Comanche, Kiowa, Lakota, Nez Perce, Miskito, and Rama (by the analyzable term *núnik kás* ‘sun/day meat;’ some of the terms have secondary associations due to the ‘cloud’-reading, see section 12). Swahili *mbingu* ‘sky’ consists of *bingu* ‘cloud’ prefixed with a noun class marker, and Biloxi *natci’ tohi’* is analyzable as ‘cloud blue’ (and denotes the ‘clear sky’ specifically). Moreover, Efik has the analyzable term *ikpa’enyöñ* /ik’pa-en’yöñ/ ‘animal.skin-sky/heaven’ used for both ‘firmament’ and ‘cloud.’ Hawaiian *aouli* ‘Firmament, sky, blue vault of heaven’ contains *ao* meaning inter alia ‘light, day, dawn, cloud’ (compare the association with ‘light’ in Swahili noted above, ‘light’ and ‘sky’ are also colexified in Guaraní), and *uli* which can refer to dark colours, including dark green and blue. A color term for ‘blue,’ as in Biloxi, also figures in other languages of North America: Upper Chehalis *tit ʔacq<sup>w</sup>éxmiʔ* is derived from *q<sup>w</sup>íx-* ‘blue,’ and Cheyenne has *otá’tavó-’omëë’e* ‘blue-realm;’ moreover, Oneida and Tuscarora colexify ‘sky’ and ‘blue.’ The association between ‘sky’ and ‘world’ in Lake Miwok corresponds to the colexification of these two meanings in Hausa (where it is archaic) and Arabela, which has a term with very wide semantic range, also including ‘earth’ and ‘hole, opening,’ which latter meaning is also colexified in Bora. And the Kashaya association with ‘rock’ is paralleled in Miskito and Hupda, which colexify the meanings (Hupda also colexifies ‘mountain’). They mirror precisely the pre-history of the Indo-European inherited term evidenced mainly by Avestan and Sanskrit evidence (Buck 1949: 52).

Four sampled languages, Yoruba, Kosarek Yale, Welsh, Hawaiian (the term here also means ‘to float,’ ‘homeless (person)’ inter alia, and denotes a particular star), and Rotuman have terms colexifying ‘sky’ with ‘air.’ In four sampled languages of Africa and Eurasia (and in some other Indo-European languages, Buck 1949: 52-53), Bakueri, Buli, Ket, and Khalkha, the word for ‘sky’ is also the name of (a) god (Bakueri also has *mmányú ya lówa* ‘up of god/sky’). ‘Sky’ is colexified with ‘top’ in general in five sampled languages, namely Efik (also colexifying ‘lift’), Hausa, Koyraboro Senni, Upper Chehalis, and Lengua, and, relatedly, in two languages, namely Hausa and Basque, ‘sky’ is colexified with ‘roof’ (the Basque term also may assume the meaning ‘canopy,’ ‘glory,’ and ‘ceiling;’ the latter meaning is also colexified in Central Yup’ik, compare also Bwe Karen *mókho* ‘sky’ and *kho* ‘top, roof’). In Buli, Khalkha, Wintu, and Rotuman, the relevant term may also assume the

meaning 'weather' and/or 'climate' (Buli also colexifies 'season, period, time' and is a "religious concept denoting the 'alter ego' or 'personal god' of an individual," while in Wintu, 'clear weather' more specifically is colexified, and this may be the relevant fact underlying the association). Guaraní and Mandarin colexify 'sky' and 'day' (see Buck 1949: 53 for Indo-European parallels), and there is a semianalyzable term in Upper Chehalis. In Buli, a term which may refer both to the 'sun' and the 'sky' is encountered, and in two languages, terms for 'sky' are based on 'sun.' These are Rama (*núnik kás* 'sun/day meat,' also colexifying 'cloud,' as noted above) and Tsafiki (*yo quido* 'sun skin'); moreover, Kapingamarangi has a semianalyzable term. Finally, Katcha and Manange colexify 'sky' with 'rain.'

Other associations include: Hausa *sama* also colexifies 'aloft' inter alia, *samaniya* dialectally also "[t]he rustle of leaves in the wind," and *gari* also 'flour, powder' as well as 'town, township,' while Khoekhoe *lātuisab* is related to *lā* 'hang (laundry) out,' 'spread out to dry,' and Swahili *anga* is derived by a zero noun-class prefix from *anga* 'light.' Baruya *sigunya* also means "fat, as of a pig," Nunggubuyu *-mala-* also means 'navel' and 'thick honey,' and *yałamara* is also the name of the "ordinary (short-horned) grasshopper" (the meaning 'sky' is rare). Rotokas *vuvui ua* is analyzable as 'transparent CLASS.NARROW.OBJECT.' The Toaripi term *kauri* is also used to refer to a tree which "has pretty sky blue flowers," while Waris *óv* also means 'to speak, for animals to make their characteristic noise.' Yir Yoront uses 'thigh' to conceptualize 'sky': *larr-kumn* is analyzable as 'place-thigh.' The term also means 'clan.' Abzakh Adyghe *we* also means 'to burst, explode' inter alia, and Basque *ortzi* also 'space' and 'storm.' Khalkha colexifies 'sky' with 'atmosphere' and Itzaj with 'to learn,' while Nuuchahnulth has a lexical suffix colexifying "in the sky" with "on a raised platform." The Pawnee term for sky, *awaahaksu*?, is analyzable as /awaahak-his-u'/ 'be.an.expanse-PERF-NOM,' and the Xicotepec de Juárez Totonac term *a'kapūn* simultaneously denotes the 'palate.' The Bora word *lévehóówa* consists of *léve* 'empty' and the classifier for doorways *-lo:gwa*. Bororo *baru* also means 'beginning,' while *waru* also means 'heat.' Cayapa *selu* also denotes a type of wave, Embera *baxã* appears to be derived from *ba* 'lightning' by means of the suffix *-xã* for covering surfaces. Guaraní *ára* is also the term for 'light' and 'time,' whereas Macaguán *bóktsebí* contains *tsebí* 'black.' Kaingang *kahnkã* also means 'family,' and the Sáliba term *mumasẽxẽ* contains *sẽxẽ* 'earth.' Tsafiki *Diósichi to* contains 'to land' (with *Diósichi* related to Span. *dios*?), Hani *aoq* also means 'yes, all right, okay,' 'to hold in mouth,' 'to sell' and is an interjection ('oh!'), Hawaiian *lani* is also the name of a very high chief and a kind of flower inter alia, and Kapingamarangi *langi* also means 'orgasm' as well as 'to commence.' Rotuman *lagi* also means 'wind' (alongside "to what purpose, wherefore"), and the Samoan term *vā-nimo-nimo* is analyzable as 'DIST-vanish-RED.'

### 53. The Smoke

Representation: 95%

Motivated: 33.9%

Thereof Analyzable: 7.0%

Thereof Colexifying: 26.9%

Thereof by Contiguity: 6.8%

Thereof by Similarity: 23.6%

Recurrent associated meanings: steam, fog, dust, cloud, fire, cigarette/tobacco, soot, spray, smell, air

Associations with other aerosols, namely, in decreasing order, with ‘steam,’ ‘fog,’ and ‘cloud’ are most frequent, not taking into account the obvious contiguity anchoring with ‘fire.’ Associations with ‘steam’ are mostly by colexification, with the exception of Piro and Tetun (*tšitšj-phy* and *ahi-suar* ‘fire-vapor’). Among the colexifying languages are Buli, Efik (by the analyzable term *nsuñ'ikañ* /n-suñ-ikañ/ ‘soft/gentle-fire’), Koyraboro Senni, Burarra, Kwoma, Mali (where the relevant term is seminanalyzable, the identifiable constituent meaning ‘liquid’), Ngaanyatjarra, Nunggubuyu, Sahu, Yir Yoront, Abzakh Adyghe, Upper Chehalis, San Mateo del Mar Huave, Central Yup’ik, Arabela, Bora, Carib, Cayapa (by the term *ñivijcha*, perhaps containing *ñi* ‘fire, flame, seed’ and *vijcha* ‘difference in height’), Guaraní, Hupda, Lengua, Miskito, Ancash and Imbabura Quechua, Tehuelche, Tsafiki, Wayampi (by the analyzable term *atasí* /atā-sí/ ‘fire-whiteness’), Fijian, Lenakel, and Samoan. Compare also Embera *kouwá*, meaning ‘hot flash’ with neuter gender alongside ‘smoke’ with masculine gender with *koúwa* ‘vapor, fume, stink.’ Buck (1949: 73) reports the pattern in Indo-European. In sixteen languages, Efik (by the analyzable term *nsuñ'ikañ* /n-suñ-ikañ/ ‘soft/gentle-fire’), Buin, Burarra, Gurindji (colexifying ‘smoke-haze’ with ‘fog’ more specifically), Yir Yoront, Abzakh Adyghe, Nez Perce (by a lexical affix), Wintu, Central Yup’ik, Hupda, Jarawara, Lengua, Maxakalí, Miskito, Rama, Wayampi (by the same analyzable term mentioned above), and Mandarin, ‘smoke’ and ‘fog’ are colexified (Buin colexifies ‘white smoke’ more specifically, in Rama there is also colexification with ‘dew’), and in nine sampled languages, Buin, Nez Perce (again by the lexical affix), Arabela, Bora, Cavineña, Cayapa (the relevant term again being *ñivijcha* mentioned above), Maxakalí, Tsafiki and Sedang (here by the analyzable term *kia hia* ‘ghost light.weight’), ‘smoke’ and ‘cloud’ are (in Buin ‘white cloud’ and ‘white smoke’ more specifically, as well as ‘to be smoking tobacco’ and ‘be affected by smoke;’ note also the similarity between Koyraboro Senni *dullu* ‘smoke,’ ‘steam,’ and *duule* ~ *dulla* ‘cloud’ as well as that the connection may be etymologically detectable for Chukchi). In some languages of the sample, more than one of the aforementioned meanings are expressed by the same term, see § 6.2.2.2. for discussion. Perhaps more surprisingly, nine sampled languages, Bezhta, Wintu (where one of the relevant term also means “cure with smoke, disinfect”), Carib, Hupda, Ancash Quechua, Fijian, Hawaiian, Sedang, and Bislama colexify ‘smoke’ with ‘dust’ (Central Yup’ik with ‘dust in air’ more specifically, and Ancash Quechua also with ‘gas;’ compare also the colexification of ‘smoke’ with “to be dust-windy” in Kiowa). By contiguity, ‘smoke’ and ‘soot’ are colexified in four sampled languages, namely Abipón, Xicotepec de Juárez Totonac, Ancash Quechua, and Toba (note that Itzaj *b'utz'* also means “blackened with soot”). Also by contiguity, ‘smoke’ and ‘smell’ are colexified in Kwoma and Cavineña (and note the evidence from Embera mentioned above). Six sampled languages colexify ‘smoke’ with ‘cigarette’ and/or ‘tobacco:’ these are Ngaanyatjarra, Abzakh Adyghe, Nuuchahnulth, Quileute, Mandarin, and Bislama (Bislama also with verbal ‘to smoke a cigarette’); the Oneida term may contain a constituent with the meaning ‘tobacco’ as well. Sedang colexifies ‘smoke’ with ‘air,’ while in White Hmong, ‘smoke’ is *pa taws* ‘air fire.’ As seen in some of the analyzable terms mentioned above, ‘fire’ is an obvious choice as a contiguity anchor for the meaning ‘smoke.’ Other complex terms of this kind include Sko *rápong* /ra-pong/ ‘fire-blow.at’ and Kosarek Yale (Obakak valley dialect) *uk solom*, perhaps analyzable



as ‘fire light.in.color.’ Cheyenne directly colexifies ‘smoke’ with ‘fire,’ and semianalyzable terms in which the identifiable constituent means ‘fire’ are found in Kwoma, Toaripi, Waris, Bwe Karen, Lenakel, Kapingamarangi, and Manange. Yir Yoront *thorrqn* also has the meaning ‘spray on waves,’ and these meanings are also colexified in Fijian and Hawaiian, which latter also colexifies ‘wisps.’ Similarly, Bislama colexifies ‘smoke’ with “spray splashed up from falling heavy rain.”

Other associations include: Efik colexifies ‘smoke, steam’ with ‘exhalation’ and ‘heat, warmth,’ as well as, from there on, “[a] good or bad influence supposed to be communicated by the heat or exhalation from the body of another.” Baruya colexifies ‘smoke’ with ‘belch, burp,’ Buin *kumogana* ‘thick white smoke’ is derived from *kumogo* ‘billow, be thick (of smoke),’ Kyaka colexifies ‘smoke, steam’ with ‘aura,’ Lavukaleve with ‘cheek,’ and Muna *ghumbo* is also used to refer to ‘many, huge numbers.’ Yir Yoront *muw* is also used as a color term for ‘gray’ (and Kiliwa *?phuuy* is also glossed as “smokey-grey”). Abzakh Adyghe *-γ°e-* has very many meanings, alongside ‘smoke, steam, fog’ also ‘dry, to make dry’ and ‘path, street,’ while Badaga *oge ~hoge* is also the term for a “burning heap of rubbish,” the ‘atmosphere,’ as well as a “gloomy state of affairs.” *Keak*, the plural of Basque *ke* ‘smoke,’ can also refer to ‘boasting, gloating.’ Biloxi *kûsidi* and *uksi’di* contain *si* ‘yellow’ and the nominalizing suffix *-di*. Yaqui *bwichia* also means ‘worm,’ and Jarawara *hasawiri/hasawiri* is also used to refer to ‘ashes.’ Yanomámi *wakě shi* is analyzable as ‘red excrement.’ Fijian *kuvu* also denotes “the foam at the front of a swiftly moving canoe,” and Kapingamarangi *huiahi* also means “to chase away, to cause to flee.” Finally, Sedang colexifies ‘smoke’ with “a blanket or shawl worn on shoulder for carrying child,” and Yay colexifies ‘smoke’ and ‘previous.’

#### 54. The Soil

Representation: 80%

Motivated: 72.4%

Thereof Analyzable: 6.4%

Thereof Colexifying: 66.2%

Thereof by Contiguity: 47.1%

Thereof by Similarity: 0.4%

Recurrent associated meanings: land/ground, dirt, world, place, dust, clay, sand, floor,

year/time, property/estate, surface, mud, field, bottom/below, day, ashes, black

For the meaning ‘soil,’ intra-domain associations are dominant. Most common is that with ‘land’ (sometimes also with ‘country’ as a political or administrative entity) and/or ‘ground,’ occurring by colexification in 67 languages, namely Efik (where the relevant term *i’sõñ* is analyzable as /i-sõñ/ ‘NMLZ-be.hard/be.firm’), Hausa (here, the relevant term also denotes a “small, red, malodorous ant,” and is also the generic name for ‘snake’ inter alia), Khoekhoe, Dongolese Nubian, Rendille, Yoruba, Baruya (where the relevant term also means ‘ladder,’ and conveys the notion of “shooting short of a target”), Buin, Burarra, Kaluli, Kwoma, Kyaka, Meyah (which also has a formally redundant analyzable term, namely *mebí efení* ‘ground/soil reflection’), Muna (also colexifying ‘island’), Ngaanyatjarra, Nunggubuyu, Rotokas, Sahu, Sko, Sentani, Tasmanian (all varieties except the Northern one), Waris, Kosarek Yale, Yir Yoront, Badaga, Basque, Chukchi, Greek, Ket, Khalkha (with

extension to many other related meanings), Laz, Nivkh (where the relevant term also means ‘edge’), Sora, Welsh, Biloxi, Highland Chontal, Haida, San Mateo del Mar Huave, Lesser Antillean Creole French, Nez Perce, Pipil, Quileute, Xicotepec de Juárez Totonac, Yaqui, Central Yup’ik (also meaning ‘village’), San Lucas Quiaviní Zapotec, Copainalá Zoque, Aguaruna, Aymara, Bororo, Carib, Cavineña, Cubeo, Embera, Huambisa, Miskito, Piro, Tsafiki, Rama, Bislama, Hawaiian, Bwe Karen, Lenakel (also colexifying ‘homeland’), Malagasy, Samoan (by the analyzable terms *‘ele’ele*, reduplicated from *‘ele* ‘compact red soil or stone, rust’ and *palapala*, reduplicated from *pala* ‘rotten’), Tetun (by the analyzable term *rai-laran* ‘earth-interior’), Vietnamese (where the relevant term also means ‘expensive’), and Yay. Lavukaleve has *ararume*, presumably containing *araru* ‘ground,’ and semianalyzable terms where one of the constituents is ‘land’ are found in Kaluli and Lenakel. Kyaka, Yir Yoront, Bezhta, Ket, Khalkha, Nivkh, Ineseño Chumash, Central Yup’ik, and Miskito furthermore colexify ‘soil’ with ‘place’ or ‘site.’

Colexification with ‘dirt’ is attested in Baruya, Buin, Gurindji, Ngaanyatjarra, Nunggubuyu, Meyah, Tasmanian (Northeastern), Waris, Yir Yoront, Ineseño Chumash, Comanche, Oneida, Pipil, Tuscarora, Wintu, San Lucas Quiaviní Zapotec, Jarawara, Hawaiian (also with ‘rubbish,’ ‘silt,’ and ‘excrement’), Lenakel (the term contains *tin* ‘land’), Rotuman, Samoan, Tetun, and Bislama (22 languages). Moreover, there is a semianalyzable term where the identifiable constituent means ‘earth, dirt’ in Lakhota.

Another recurring pattern is the association with ‘dust,’ occurring in eight of the sampled languages, Efik, Yoruba, Dadibi, Ngaanyatjarra, Northeastern and Southeastern Tasmanian, Oneida, Wintu, and Toba (and evidenced by cognates from Old English and Gothic, one meaning ‘soil,’ the other ‘dust,’ Buck 1949: 18). Seven further sampled languages colexify ‘soil’ with ‘clay,’ namely Buli, Efik, Baruya, Bezhta, Japanese, Hupda (also colexifying ‘Tuyuca people’), and Bislama, and, similarly, in Koyraboro Senni, San Mateo del Mar Huave, and Samoan, the same term is used for both ‘soil’ and ‘mud.’ Bezhta and Cubeo colexify ‘soil’ with ‘field,’ and Lesser Antillean Creole French, San Lucas Quiaviní Zapotec, and Mandarin with ‘property, estate.’

In six sampled languages, Efik (by the same analyzable term mentioned above), Basque, Guaraní, Macaguán, Imbabura Quechua, and Tetun, the term for ‘soil’ is also used for the ‘floor’ (e.g. of a house), and in another six, Buli, Efik, Rendille, Ngaanyatjarra, Tehuelche and Rotuman, ‘soil’ and ‘sand’ are colexified (the Buli term is semianalyzable, containing *tain* ‘stone, pebble,’ and the Tehuelche term also means ‘sand dune’). Kapingamarangi has the corresponding analyzable term *gelegele luuli* ‘sand black’ (compare also Tetun *rai-metan* ‘earth-black’). Oneida and Wintu colexify ‘soil’ also with ‘ashes;’ the relevant Wintu term *bukul* is related to *buk* ‘dark.’

There are also associations which move to more abstract spatial relations. For instance, the Efik term mentioned above, as well as an unanalyzable Hausa term also mean ‘bottom, below’ (see Heine and Kuteva 2002: 121 generally and Buck 1949: 18 for a parallel from Latin; note also that Embera *udáa* has the adjectival meaning ‘down’ alongside nominal ‘soil, ground’), and also in Efik, as well as in Abzakh Adyghe and Bora, terms bearing an association with ‘surface’ in general occur: Abzakh Adyghe *s’əg°-šheš°ə* is analyzable as ‘earth-surface,’ and Bora *íñúji hallu* consists of *íñú* ‘earth’ suffixed with the classifier for

disc-like objects *-ji* and *hallu* ‘top, outside part.’ In Efik, Rendille, Abzakh Adyghe, Basque, Ineseño Chumash, Quileute, San Lucas Quiaviní Zapotec, Aguaruna, Carib, Guaraní, Huambisa, and Ancash Quechua, as in other European languages, terms for ‘soil’ may also refer to the entire ‘world.’

Even more interesting are extensions into the temporal domain, that is, terms for ‘soil’ than can also refer to a ‘year’ and/or ‘time’ in general. Kyaka, Ket, Highland Chontal, Ineseño Chumash, and Ancash Quechua are languages in the sample with terms that behave like this, while Kyaka and Yir Yoront colexify ‘soil’ with ‘day’ (in fact, *larr*, the relevant Yir Yoront term has a very wide semantic range); Holmer (1966) also reports similar patterns of colexification for languages of Oceania. Kyaka furthermore colexifies “event, opportunity, chance” as well as ‘weather.’

Other associations are: Efik *isõñ* colexifies ‘soil’ with “hatch or trap door,” *umabu* colexifies ‘soil’ with ‘mould,’ and *n'tan* is also the name of a plant with flowers that eject a dusty substance when being touched. Toaripi *mea* colexifies ‘soil’ with ‘wind, weather’ inter alia, while Sahu *tana'a unisi* “dry, infertile soil on a ridge” is analyzable as ‘earth/ground/land shin/shin.bone.’ Kosarek Yale *soko*, dialectally used with the meaning ‘earth, soil’ also denotes a ‘special type of tie rod’ without dialectal restrictions. Abzakh Adyghe *ś'ə* also means ‘to produce, to construct’ inter alia, and *ś'əg*, a complex term of the redundant kind, can also refer to the ‘ground floor.’ Badaga *parava* means “bothered, concerned, troubled” and is also the name of a specific kind of soil, and Khalkha *kərysın* ~ *kərydesın* also means ‘crust, peel, rind.’ Welsh *gweryd* is also used as a term for ‘grave.’ The root *-ir-* yielding Tuscarora *à-wi?r* ‘soil’ can also refer to a “bit, grain, particle” and ‘small piece,’ Kaingang *ga* also means ‘louse, worm,’ and Piro *tšì-xi* appears to be a diminutive of *tšì* ‘fire, firewood.’ Rama colexifies ‘earth, ground’ with ‘going,’ Fijian with ‘cluster, shoal, swarm,’ and Bwe Karen *ha* also means ‘hole in the ground, pit’ inter alia. Lenakel *nimitik* ‘red or reddish soil’ might contain *nimit* ‘mud, swamp,’ Hawaiian *lepo* figuratively can also refer to ‘common people,’ Samoan ‘*ele'ele*’ (see above for morphological analysis) also means ‘blood’ and ‘menses’ in polite usage, while Tetun *rai-metan* ‘earth-black’ colexifies ‘fertilizer.’

### 55. *The Spark*

Representation: 54%

Motivated: 39.1%

Thereof Analyzable: 27.2%

Thereof Colexifying: 13.1%

Thereof by Contiguity: 10.0%

Thereof by Similarity: 23.1%

Recurrent associated meanings: fire, lightning, embers, flame, burst/explode, light, particle, sparkle, firefly, flower, star, grain

There are a number of lexico-semantic ties pertaining to the meaning ‘spark,’ but none of them is particularly frequent. Basque, Khalkha, Haida, and Tuscarora colexify ‘spark’ with ‘embers’ (Tuscarora in addition colexifies ‘candle,’ ‘flash of light,’ ‘light,’ ‘lamp,’ and ‘taper’). In three languages, the term for ‘spark’ is associated lexically with verbs meaning ‘to burst, explode.’ These are Nivkh (*p'ryrk t'uyr* ‘burst fire’), Nez Perce (*taxlicá'sa*, containing

*tax* ‘explode’ and *lićé* ‘be directed at’), and Fijian (*lidi ni buka* ‘burst/explode POSS fire/firewood’). The same word is used for ‘spark’ and ‘flame’ in four sampled languages, Abzakh Adyghe, San Mateo del Mar Huave, Bororo, and Chayahuita (the Abzakh Adyghe and Chayahuita terms in fact are complex containing the words for ‘tongue,’ since this is a very frequent pattern in terms for ‘flame’ as described in section 22, it seems reasonable to assume that this is indeed the primary meaning in these languages). Tuscarora and Central Yup’ik have associations between ‘spark’ and ‘light’ (Central Yup’ik by the term *kenurraq*, which is perhaps analyzable as /keneq-rraq/ ‘fire-little.bit,’ the meaning ‘spark’ is attested only for the dialect of Norton Sound, in other dialects it means ‘light’ or ‘lamp’). Relatedly, in Khoekhoe, *nanib* ~ *nanis* is derived from the verb *nani* ‘to twinkle, flicker, gleam, burn slowly,’ and there is a semianalyzable term in Kaingang. Hawaiian and Rotuman have complex terms in which one constituent means ‘particle:’ *huna-ahi* ‘particle/speck/crumb/grain-fire’ (also denoting ‘live cinder’) and *momoe ne rāhi* ‘fine.particles ART.PL fire.’ Also in two languages of the sample, an association with the meaning ‘to sparkle’ is found, by colexification in Bororo and by the term *yantsáji* /*yáants-ji*/ ‘sparkle-POSS’ in Aguaruna. Contiguity-based associations by morphologically complex terms with ‘fire’ making reference to the fact that sparks are, in Cognitive Linguistics parlance, a figure emerging from a larger ground structure (the fire) are Kashaya *?oho cuh<sup>h</sup>u?-cuh<sup>h</sup>u?w*/ ‘fire round.object-pieces.come.off.bigger.object-RED-ABS’ and Wintu *p<sup>h</sup>o-h dil-ma* ‘fire drop/fall/alight-??’ (this term itself is glossed as “Sparks are flying. He dropped the fire”). There is also at least one language which directly colexifies ‘fire’ with ‘spark,’ namely Chukchi; the association may also be present in Northeastern Tasmanian.

Recurrent metaphor-based associations are also found. Abzakh Adyghe, Badaga, Abipón, Bora, Ancash Quechua, and Lesser Antillean Creole French colexify ‘spark’ with ‘lightning.’ Berik has *tokwa es*, presumably analyzable as /tokwa ese/ ‘fire flower’ and Tetun *ahi-fuhan*, also analyzable as ‘fire-flower.’ In two sampled languages, the word for ‘spark’ contains that for ‘star’: Dadibi *sia hó* ‘fire star’ and Bislama *sta blong faea* ‘star POSS fire.’ In Baruya, ‘spark’ is *dí’nyaala* /*dika-nyaala*/ ‘fire-firefly,’ Bezhta directly colexifies the relevant meanings, while Huambisa has the semianalyzable term *yantsari*, for which compare *yantsa* ‘firefly.’ Miskito has *pauta yuya* ‘fire grain,’ and Hawaiian *huna-ahi* ‘particle/speck/crumb/grain-fire.’ Other metaphorical associations in which ‘fire’ acts as a contiguity anchor include: Yoruba *owó-iná* ‘hand-fire,’ Toaripi *a-e*, perhaps analyzable as ‘fire-faeces,’ Ket *bógdes*, analyzable as /bo<sup>7</sup>k-dēs/ ‘fire-eye’ (note that Haida *sráahld* is derived from a verb meaning ‘to glance’), Xicotepec de Juárez Totonac *kosa macscut* ‘jump/get.up fire.’ Semianalyzable terms in which ‘fire’ figures are found in Efik, Toaripi, Kosarek Yale, Carrier, San Lucas Quiavini Zapotec, and Arabela.

Other unsystematic associations include the following: Muna *wara*, when used verbally, also means ‘to drizzle, to drip,’ and Yoruba *èta* also means ‘splash’ as well as ‘root, tuber.’ Ngaanyatjarra *tii* also means “healed tissue” as well as, by English influence, ‘tea,’ while Nunggubuyu *-rarwadawada-* may be related to *rār* ‘burnt-out grassland or light bushland’ and *=w<sub>2</sub>ada-* ‘to snap, to break suddenly.’ The second constituent in the Yir Yoront term *thumlilqli* resembles *lilq*, which means ‘alone, by oneself’ (*thum* is ‘fire’).

Basque *pindar* also colloquially means “spunk, pep,” and *txinpart* is also used with the meaning ‘livewire.’ Khalkha *cindara* contains or is otherwise related to *cindar*, a respectful term for ‘remains, corpse’ and also means “white ashes on dying coals.” Coy also denotes ‘hot ashes,’ as well as, figuratively, “splendor, grandeur, glory; energy, spirit.” Welsh has *tamaid o beth llosg* ‘piece of something burning,’ and Blackfoot *ipáísstsitiimi’kaa* ‘to throw a spark’ contains *mi’k* ‘red.’ The denotational range of the Itzaj term *se’es* covers also “sawdust, bit, small chip, crumb, confetti” as well as “finely cut, fine.” The Kiliwa term contains an element meaning ‘earth.’ Wappo *pétsi?* also means “snap like burning wood,” and Copainalá Zoque colexifies ‘light thunder.’ Arabela colexifies ‘piece of lit coal,’ and Bororo *beri* is also used with the meanings ‘abundance’ and ‘arrogance,’ while Embera *adyizúa* colexifies ‘brilliant’ and ‘bright.’ Great Andamanese *châpal’igbêra* contains *châpa* ‘firewood’ and *bêra* ‘sweepings,’ and Hani *miqseil* might contain *seil* ‘louse’ (the term also denotes a ‘woman dissatisfied with her marriage’). Kapingamarangi *madagologolo* is related to *kolo* ‘to drill, twist as a knob.’ Samoan *sipaka* (< Engl. *spark*) also denotes a ‘spark plug.’

#### 56. *The Spring*

Representation: 87%

Motivated: 46.3%

Thereof Analyzable: 35.5%

Thereof Colexifying: 12.0%

Thereof by Contiguity: 27.8%

Thereof by Similarity: 12.6%

Recurrent associated meanings: water, river, hole, eye, come out, lake/pond,  
jump, head, puddle, boil, dig/dug, fetch water, headwaters

Motivated terms for ‘spring’ (or ‘well,’ which was accepted as a proxy) are frequently morphologically complex, with one of the constituents being ‘water.’ However, also for the meaning of the second constituent involved, there are a number of recurrent patterns. In twelve sampled languages, this is ‘hole,’ as in Bakueri *ewondí yá málíwá* ‘hole of water.’ Alongside Bakueri, words with this structure are attested for Berik, Toaripi, Haida, Yuki, Bora, Bororo, Guaraní, Hupda (where the word for ‘hole’ also means ‘house’), Maxakali, Rama, Yanomámi (where an additional element meaning ‘point’ is present), and White Hmong (where there may be a further element meaning ‘issue forth’ present). In addition, Embera colexifies a general term for ‘hole, pit’ with ‘spring’ and in Hani, ‘village well’ is *lolhovq*, with *lol* being the classifier for rivers and *hovq*, meaning ‘to fetch water’ (for which association in turn compare Chukchi *ajmā-n* ‘fetch.water-LOC’) as well as being a classifier for pits and holes (see also Buck 1949: 45 for this semantic connection in Indo-European, evidenced in Greek and Armenian). Somewhat similarly, Cubeo has *jiacarã-cobe*, probably containing *jiacacæ* ‘be.aquatic’ and a classifier for hole-like objects. The second most frequent pattern is metaphorical in nature, with the second constituent being ‘eye,’ as in Meyah *mei eitíj* ‘water eye’ (see also § 6.2.3.1. for discussion of ‘eye’-metaphors). Other languages with such terms are Buli, Kyaka, Sahu, San Mateo del Mar Huave, Bislama, Fijian, and Tetun (note also Welsh *llygad ffynnon* ‘eye well/fountain’ and similar redundant terms in Khoekhoe, Muna, Ancash Quechua, and Hawaiian). Hausa, Dongolese Nubian, Burarra, Ancash Quechua, and Samoan colexify ‘eye’ with ‘spring,’ due to the pervasive-

ness of metaphors on the basis of ‘eye’ cross-linguistically (see Buck 1949: 44 for semantic development from ‘eye’ to ‘spring’ in Armenian), many of the languages have further meanings colexified. The association with ‘eye’ is also common in Semitic languages (e.g. Segert 1991: 1432). Another metaphorical pattern is constituted by the transfer from ‘head’ to ‘spring,’ as in Yanomámi *u he* ‘liquid/river head.’ This association is also found in Mbum, Sora (where an additional element meaning ‘hill’ is present and the relevant term denotes a “spring of water on the hill”), Malagasy, and Tetun; Miskito colexifies ‘head of plant’ and ‘spring,’ and a semianalyzable term of this kind, colexifying ‘source’ (also in the sense of ‘source of information’) and ‘origin’ is also found in Basque; also note Yoruba *orí-sun* ‘head-spring.or.fountain.’ In four sampled languages, Mbum, Yoruba, Ancash Quechua, and Rotuman, terms for ‘spring’ are encountered which contain a verb meaning ‘to jump’ or to ‘spring,’ for instance Mbum *hvíŋà-mbìì* ‘spring.out water’ for a ‘fountain;’ the relevant Fijian term is reduplicated from a verb meaning ‘for water to spring up.’ Samoan colexifies verbal ‘jump, leap’ with ‘boil’ and nominal ‘spring, source’ (for the association with boiling, note that Central Yup’ik *qalla-neq* colexifies ‘spring’ with ‘eddy’ and is analyzable as ‘be.boiling-thing.that.results.from’ and that the Swahili term *chemchemi* is derived by reduplication from the verb *chemka* ‘to boil’). Somewhat similarly, in eight sampled languages, One, Japanese, Nivkh, Sora, Kiowa, Bora, Chayahuita, and Takia (in Japanese and Bora, there are uncertainties as to the analysis), there are at times quite complex terms for ‘spring’ revolving around verbs meaning ‘to come out,’ ‘to go out,’ or ‘to exit,’ such as One *folá suwe* ‘water/river come.out;’ note also that in Kiliwa *?matcpam*, the sequence *-cpam* (*mat* is ‘earth’) might derive diachronically from *\*c+paa* ‘come out.’ Buin colexifies ‘spring,’ ‘to emerge’ and other meanings directly, and Welsh *tarddiad* is derived from *tarddu* “to spring, to sprout, to derive from, to issue.” In Biloxi, perhaps Cheyenne, Lakshota, and Tehuelche, constituents meaning ‘to dig (out)’ or ‘dug’ figure, e.g. in Lakshota one of the terms for ‘spring’ is *mničłápi* /mničłápi/ ‘water-dug’ (Cheyenne colexifies ‘spring with ‘water pump’ and ‘windmill’); Rotokas and Copainalá Zoque have semianalyzable terms with such structure.

Twelve sampled languages, Bakuéri, Buli, Efik, Ngambay, Muna, Sko, Badaga, Kashaya, Pipil, Maxakalí (by the complex term *kōnāgkox*, analyzable as /kōnā’āg-kox/ ‘water-hole’), Bwe Karen, and Kapingamarangi directly colexify ‘spring’ or ‘well’ with ‘river, stream’ by contiguity (Ngambay also with ‘waves;’ the connection is evidenced by a cognate set between Latvian and Sanskrit in Indo-European, Buck 1949: 45). Buli, Wintu, Yuki, Cayapa, Kaingang, and Maxakalí colexify ‘spring’ with ‘pool, lake’ and/or ‘pond’ (again by the analyzable term already mentioned). Similarly, in Sko, the ‘spring’ is called *pa-í* ‘water/river-pool.’ In Basque, Oneida, Pipil, and Hawaiian, ‘spring, well’ is colexified with ‘puddle.’ Due to colexification of ‘water,’ ‘river,’ and ‘lake,’ the association with ‘lake’ is also found in Berik by an analyzable term, and there is a semianalyzable term in Quileute.

This brings the discussion back full circle to the frequent role of ‘water’ to act as contiguity anchor in complex terms, and the different patterns of colexifications of ‘water’ with different bodies of water (see also § 6.2.2.5.). Ngambay, Muna, Yir Yoront, Comanche, and Kashaya directly colexify ‘(spring) water’ and ‘spring’ (and sometimes also other bod-

ies of water, further, Ngambay also colexifies ‘waves’ and Badaga also ‘dale,’ ‘flat land’ and ‘riverside’).

Noni *fjoo* ‘spring’ consists of *joo* ‘water’ and the noun class prefix *fi-*, and Lavukaleve *lafio* ‘spring’ is connected to *lafi* ‘water.’ Bwe Karen has *chí-bú* ‘water-in.’ However, alongside those already mentioned, there are also a relatively large number of language-specific conceptualizations realized by morphologically complex terms on the basis of ‘water.’ These are: Noni *joo yi caan* ‘water REL small,’ Kaluli *ho:n sí* ‘water tip’ (compare the possible etymology of Ket *tájlop* < \**taj-ül-?qōp* ‘cold-water-tip’), Kyaka *ipwua renye* ‘water source’ (with *renye* also having other readings), Meyah *mei ofog* ‘water round,’ Sahu ‘*banyo ma utu’u* ‘water POSS root,’ Abzakh Adyghe *psə-łaq<sup>o</sup>e* ‘water-foot/stem,’ Ket *aqtul*, analyzable as /aqtə-ül/ ‘good-water,’ Carrier *thaḳezḳēt*, containing *tha* ‘water’ and *ḳēt* ‘place,’ Oneida *kahnekóni?* /ka-hnek-No-?/ ‘NEUT.AGENT-liquid/liquor-be.in.water/cook.in.water-STAT,’ Pipil (Santo Domingo de Guzmán dialect) *pu:ní a:t* ‘be.born water-ABS,’ Santiago Mexquititlan Otomí *pothe* /poho-dehe/ ‘well.up-water,’ Wintu *popil-mem* ‘summer-water/river’ (also denoting a “spring that dries in the winter”), Yuki *pil-?u?uk* ‘snow-water,’ Guaraní *y-vu* ‘water swollen,’ Wayampi *iapi* /i-?api/ ‘water/river-source,’ colexifying ‘valley’ and ‘dew,’ Bislama *maot blong wota* ‘mouth POSS water,’ and Hawaiian *kumu wai* ‘foundation/base water’ and *wai hū* ‘swell water’ (meaning “gushing spring, overflowing water”). There are semianalyzable terms with ‘water’ in Kosarek Yale, Comanche, Quileute, Yana, Yuki, Bislama, and Kapingamarangi. Baruya and Kosarek Yale colexify ‘spring’ with ‘headwaters,’ and similarly, Sedang colexifies ‘upstream.’

Other associations include: the Khoekhoe terms *laub* and *laus* contain the root *lau* “trickle, purl, run/flow gently,” and Rendille *wór* also means ‘news.’ Anggor *fe amongo* seems to contain *amongo* ‘sibling,’ Basque colexifies ‘well’ and ‘puddle,’ while Sahu *gogonyo’o* contains *onyo’o* ‘to draw water.’ Khalkha *bulay* also means “[h]aving white spots, partly white...,” and Welsh *ffynhonnell* is derived from *ffynhonni* ‘to well, to gush.’ Ineseño Chumash ‘*aqmilimu*’ is derived from the verb ‘*aqmil*’ ‘to drink,’ while Itzaj colexifies ‘spring’ with ‘splash’ and Lake Miwok *?óla* is also a kinship term. The Pawnee term *kicaahkatakus* is analyzable as /kic-haahka-ta-kus/ ‘be.liquid-be.attached-suspended-be.sitting,’ Santiago Mexquititlan Otomí colexifies ‘spring’ with ‘black,’ and the Tuscarora term *uča?tuhsa?kyéha?* is based on the verb root -*ča?tuhsT*- ‘be cool.’ Guaraní *ykuayvu* also denotes an ‘underground watercourse.’ Hawaiian *hāpuna* colexifies ‘spring’ with ‘coral,’ ‘lime’ and other things, and *māpuna* “bubbling spring” with “froth, as of a rough sea” and “surging of emotions.” Mandarin colexifies ‘spring’ and ‘neck’ (the relevant lexical items were still distinct in Early Middle Chinese though), and Rotuman colexifies ‘water source’ with ‘medicine,’ ‘cask,’ and “stew or hash made of meat or fish.”

### 57. The Star

Representation: 97%

Motivated: 21.2%

Thereof Analyzable: 5.2%

Thereof by Contiguity: 2.7%

Thereof Colexifying: 18.1%

Thereof by Similarity: 19.2%

Recurrent associated meanings: planet, firefly, starfish, meteoroid, moon, constellation, blaze, shine/sparkle/blink, asterisk, badge of rank, dot/spot, fire

‘Star’ is a meaning expressed in many languages by an unanalyzable, monomorphemic word. Semantic associations by colexification are also relatively rare. Muna, Nunggubuyu, Rotokas, Toaripi, Tuscarora, Bororo, Wayampi, Yanomámi, Fijian (where the relevant term *kalokalo* also denotes the flower ‘aster,’ the English name of which incidentally itself goes back to the Ancient Greek word for ‘star;’ *kalo* itself is ‘pull bowstring, discharge gun’), and Sedang colexify ‘star’ with ‘planet’ (Sora and Toaripi also with ‘comet’ or ‘meteor;’ moreover, in Khoekhoe, the same term suffixed with different nominal designants yields the meanings ‘star’ and ‘comet’ respectively, while Bislama *sta* is glossed as “any heavenly body (e.g. moon, star, meteorite)”). The second most common association, found in Buin, Muna (where the term *kolipopo* ~ *ngkolipopo* may be related to *popo* “evil spirit which looks like a flashlight attacking people”), One, Waris, Kosarek Yale, San Lucas Quiaviní Zapotec, Bora, Cavineña, and Yanomámi, is colexification of ‘star’ with ‘firefly’ by a metaphorical transfer based on perceptual similarity; the relevant Yanomámi term is semianalyzable containing an element meaning ‘round fruit’ and also denotes an unidentified species of *oruga*.

Otherwise, there is scattered evidence for lexico-semantic ties between ‘star’ and the two large heavenly bodies, the ‘sun’ and the ‘moon.’ Burarra is the only language in the sample which colexifies ‘sun’ and ‘star’ (and concomitantly, a type of shellfish similar in appearance to a star, several species of sea urchin, and further meanings associated with ‘sun,’ such as ‘watch,’ see section 60), though note that Burarra also features an unrelated monomorphemic term for ‘star’ specifically. Semianalyzable terms for ‘star’ in which one constituent appears to be ‘sun’ are found in Maxakalí and White Hmong, and in addition, Hupda *wædhəm’ǎh* might be analyzable as */wædhə-məh/* ‘sun/moon-small,’ although this is not entirely straightforward. The source remarks that the words for ‘star’ are also based on a word for ‘sun, moon’ in languages of the neighboring Tucanoan family. Ties with ‘moon’ specifically also exist. Next to the case of Bislama already mentioned, Abipón is the only language in the sample colexifying the two by the term *eergRaik*, derived from *eerg-* ‘to burn, sparkle.’ In the two sampled Tupi-Guaraní languages, Guaraní and Wayampi, the word for ‘star’ is analyzable as ‘fire-moon’ (*jasyrata* */jasy-tata/* and *yai-tata* respectively; Guaraní also has the variant *yvágara* */yvaga-tata/* ‘sky-fire’), and in addition, there are a number of languages in the sample where a diachronic association may exist. These are suspiciously concentrated in Africa: the Buli word for ‘star,’ *chingmarik*, might contain *chiik* ‘moon’ and *ngmari* ‘take from’ (this analysis is marked as questionable in the source). Likewise the analysis of Efik *ntan’tafiōñ* is dubious: it might be a complex term consisting of *n’tan* ‘earth, dust’ and *ō’fiōñ* ‘moon.’ Koyraboro Senni is yet another African language with a term for ‘star’ that is apparently related, at least in a diachronic sense, to other lexical elements: *handarey* resembles both *handi* ‘day’ and *handu* ‘moon, month.’ For Dongolese Nubian *wíss(i)*, Armbruster (1965) suggests an etymology connecting the term to a word for ‘moon’ plus a diminutive suffix; note also the similarity between Rendille *yeyyehím* ‘star’ and *yéyyah* ‘moon.’



Moreover, Carib, Tehuelche, and Wayampi use the same term for a 'star' and a 'constellation of stars,' and Toaripi, Sora, and Bislama employ a single term for both 'star' and 'shooting star, meteor.'

Another class of terms for 'star' are those derived from verbs meaning 'to shine,' 'to sparkle,' 'to blink' etc. (see Buck 1949: 56 for the association with 'shine' in Sanskrit). As already mentioned, Abipón *eergRa-ik* is derived from a verb meaning 'to sparkle' and similarly, in Khoekhoe, the word for 'star,' *lgami-ro-s*, is analyzable as 'blink-DIM-3SG.FEM' (a variant is *lhom-lgami-ro-s* 'sky-blink-DIM-3SG.FEM'). In Hani, the 'star' is called *aqgeel alsiq* or *aqgeelsiq*, with *geel* meaning 'to shine' and *alsiq* 'fruit'; *siq* also acts as a classifier for round things, inter alia (for 'fruit,' compare Austin et al. 1976: 61, table 2 for evidence from Arabana and Wanganuru), and Piro *katahiri* apparently contains *kata* 'shinging, blazing.' Three sampled languages, Hausa, Basque, and Kildin Saami, colexify 'star' with 'blaze'; the Basque term also may refer to a 'cataract,' 'asterisk' (an association it shares with Lesser Antillean Creole French), 'dream,' and a 'star' in the sense of a celebrity as well as someone 'dear, beloved.'

Other recurrent associations include: In four languages of Oceania, Buin, Lavukaleve, Rotokas, and Bislama, 'star' and 'starfish' are designated by the same term. In two sampled languages of Eurasia, Basque and Khalkha, 'star' is extended to 'badge of rank' (a pattern also found in English, German, and presumably other languages of the region). Kyaka and Oneida colexify 'star' with 'dot' and/or 'spot' (Oneida also with 'print').

Still other associations include: Buin *kaipa* may also refer to "anything star-shaped," Gurindji *kiki* also to a type of ornament, while Meyah *motúr* is semianalyzable: it contains *motú* 'night.' One *leila* also denotes a river frog, Sko *ha* also means 'bag' and 'walk,' Kosarek Yale has a semianalyzable term containing an element meaning 'sky, air,' Waris *pai* is also an interjection ("my!"), Badaga *mi:nu* also means 'fish, shellfish,' Bezhta *cā* also 'salt,' and Kolya Yukaghir *jurgud'e:jə ~ jurgud'e:jjə*, colexifying 'awl,' contains *jurgu*: 'slot, hole' (Tundra Yukaghir has *payad'iid-ekuu* 'drill-hole,' Nikolaeva 2006: 340). Ineseño Chumash *'aqiwo* also means 'snail.' The Haida term *k'a7ihldaa* is analyzable as /k'a-7ahlda/ 'tiny.object-glance.at,' Itzaj colexifies 'star' with 'black,' and the Kashaya term *q<sup>h</sup>a'moš* contains *moš* 'sour.' The first element may be etymologically related to *q<sup>h</sup>a'ʔa* 'nightlong' or *q<sup>h</sup>a'aw* 'morning.' Tuscarora *uʔnihsə'reh* also means 'navel.' Wintu *luyuw* is related to *lu* 'to stab' and also denotes 'porcupine, porcupine needles.' Arabela *rijia* also means 'earthquake.' Chayahuita *tayora* probably contains the classifier *-ra* for 'small things,' and Lengua *yoa* also denotes a 'pebble.' Miskito *karma* also means 'throat' and 'origin.' Rama *piúp* contains *up* 'eye,' and Yanomámi *kurikayari* might be related to *kurikaya*, a term for a parrot species. Bwe Karen colexifies 'star' with 'to run' inter alia, and Rotuman *sina* also means 'light, lamp' inter alia (for this term the lexicographer remarks that it might be restricted with the meaning 'star' to a single fixed expression). The Yay term for 'star,' *daaw<sup>i</sup> di<sup>i</sup>*, might contain *daaw<sup>i</sup>* 'to stir,' and Vietnamese *sao* also means 'how.' Lesser Antillean Creole French *étwal* also means 'destiny.'

58. *The Steam*

Representation: 70%

Motivated: 53.9%

Thereof Analyzable: 11.3%

Thereof Colexifying: 42.7%

Thereof by Contiguity: 17.5%

Thereof by Similarity: 29.4%

Recurrent associated meanings: smoke, fog/mist, breath/exhalation, heat/hot, smell, cloud, water, dust, air, gas, boil, spray, sweat, fire

The meaning ‘steam’ (for which ‘vapor’ was accepted as a proxy) is frequently associated with other dispersions of particles in the air, such as ‘smoke,’ ‘cloud,’ and ‘fog.’ 31 languages colexify ‘steam’ with ‘smoke;’ these are Buli, Efik, Koyraboro Senni, Burarra, Kwoma, Mali (the relevant term *chulēski* contains *chulēs* ‘liquid’), Ngaanyatjarra, Nunggubuyu, Sahu, Yir Yoront, Abzakh Adyghe, Upper Chehalis, San Mateo del Mar Huave, Central Yup’ik, Arabela, Bora, Carib, Cayapa (by the term *ñivijcha*, perhaps containing *ñi* ‘fire, flame, seed’ and *vijcha* ‘difference in height’), Guaraní, Hupda, Lengua, Miskito, Ancash and Imbabura Quechua, Tehuelche, Tsafiki, Wayampi, Fijian, Lenakel, and Samoan. Morphologically complex terms are also found: in Highland Chontal, ‘steam’ is *liguxís gajah* /*liguxís lajah*/ ‘smoke water,’ and in Santiago Mexquititlan Otomí *bipa* is analyzable as /*bifi-pa*/ ‘smoke-heat.’ In Hawaiian, one term for ‘steam’ is *uahi wai*, where *uahi* ~ *uwahi* means ‘smoke, spray’ and *wai* ‘water, liquid’ (compare also the colexification of ‘steam’ with ‘spray from waves’ in Yir Yoront and with ‘spray’ generally in Fijian). Similarly, Lake Miwok *šimúuṭi* also means ‘for smoke to emerge.’

In addition, in two of the languages with terms colexifying ‘smoke’ with ‘steam,’ they are morphologically complex, and their internal structure suggests that ‘smoke’ is the dominant meaning. Efik *nsuñ’ikañ* contains *suñ* ‘soft gentle’ and *ikañ* ‘fire,’ and in Wayampi, *atāsī* is analyzable as /*ata-sī*/ ‘fire-whiteness;’ moreover, the relevant Kwoma term is semianalyzable, with the identifiable constituent meaning ‘fire.’ However, the direction of mapping as revealed by the evidence from complex terms is not entirely unidirectional, as evidence from Tetun shows (see section 53).

Sixteen languages in the sample, namely Efik, Anggor, Burarra, Kyaka, Yir Yoront, Abzakh Adyghe, Greek, Khalkha, Welsh, Wintu, Arabela, Hupda, Lengua, Miskito, Hawaiian, and Rotuman colexify ‘steam’ with ‘fog, mist’ (see Buck 1949: 66 for the connection with ‘vapor’ in Indo-European), and five, Anggor, Arabela, Bora, Cayapa (by the analyzable term mentioned above), and Tsafiki colexify ‘steam’ with ‘cloud’ (note also the similarity between Koyraboro Senni *duule* ~ *duula* ‘cloud’ and *dullu* ‘smoke, steam’). Moreover, four sampled languages, Carib, Hupda, Ancash Quechua, and Fijian colexify ‘steam’ with ‘dust,’ and in Wintu, Ancash Quechua, and Hani, the same term is used for ‘steam’ and ‘gas’ (the relevant Wintu root *xO*s also yields *xoso* ‘lungs’). As is always the case with the associations pertaining to aerosols, it should be borne in mind that some sampled languages use the same term for not only two, but sometimes three or even four of these meanings (see also § 6.2.2.2.).

An association that is particular to the meaning ‘steam’ is that with ‘breath, exhalation,’ occurring by colexification in Bakueri, Efik, Buin (where the term also denotes the

'windpipe, trachea' as well as "Puffing, breathlessness; airiness"), Kosarek Yale, Bezhta, Khalkha, San Lucas Quiaviní Zapotec, Ancash Quechua, Fijian, Hani, and Sedang (inter alia), and by the analyzable term *shdl-hǎ'-t'əu-gyH* 'be.hot-breath-water-NOUN.POSTFIX' in Kiowa. As in this language, analyzable terms for 'steam' frequently contain an element referring to 'heat' (and Efik, Yoruba, Itzaj, and Pawnee, by the analyzable term *awiriitu'* /*awirit-u'* 'be.hot-NOM,' colexify the meanings directly, while Bislama colexifies 'steam' with "radiated heat from sea" more specifically). Santiago Mexquititlan Otomí has, as already mentioned above, *bipa*, analyzable as /*bifi-pa*/ 'smoke-heat,' Yanomámi has *u heōshi* 'liquid hot,' and Takia *you wanna-n san ur* 'water hot-3SG its air' (for the association with 'air,' note also Vietnamese *hơi nước* 'air water' and the colexification of these meanings in Bakueri and Ket alongside the colexification with 'atmosphere' in Khalkha). A semianalyzable term where one constituent can be identified to mean 'hot' is found in Rotuman. Completely unrelated patterns are the colexification of 'steam' with 'smell, odor' and/or 'fragrance (from cooking),' which is found in Ngambay, Burarra (here the relevant term *-jinyja* is related to the verb *jinyja*, meaning 'be standing, be vertical' inter alia), Gurindji, Kwoma, Basque (where the term also denotes 'perfume' specifically), Bezhta, Tuscarora (by an analyzable term containing the roots *-hsəri-* 'savor' and *-ur-* 'to cover'), and Yay. Colexification of 'steam' and '(to) sweat' is found in Santiago Mexquititlan Otomí (by the analyzable term *xani-dehe* 'to sprinkle-water') and White Hmong. As has emerged from the discussion so far, 'water,' associated by contiguity with 'steam,' frequently figures in analyzable terms. Another such term not yet mentioned is Santiago Mexquititlan Otomí *hœ-dehe* 'to.fall-water,' and semianalyzable terms where one of the constituents is 'water' are also found in Kiliwa and Sko alongside Japanese, where a term for 'steam' contains *yu* 'hot water' (see Goddard 2001 on the lexical distinction between hot and cold water in Japanese). Abipón has *l-apa-Ra*, analyzable as 'POSS.INDEF/3SG-boil-ABSTR' and colexifying 'foam,' Great Andamanese *bôag* has a verbal reading as 'to boil,' and a verb colexifying 'to steam' with 'to boil' is present in the relevant term in Oneida.

Other associations include: Efik *uye* also denotes "[a] good or bad influence supposed to be communicated by the heat or exhalation from the body of another," and Muna *oho* also means 'to feed' as well as 'close lid, cover.' Kosarek Yale *iba* metaphorically also means 'fury, passion.' Abzakh Adyghe *-y'e-* also means 'to dry, to make dry' and 'path, street' among many other things, Badaga *a:vi* also means 'yawn, yawning,' as well as 'spirit, soul' and related notions, Khalkha *ayur* metaphorically also means 'anger,' while Welsh *anwedd* also means 'enormous' and other things. Cahuilla *múlulš* is derived from the verb *-múlul-* 'to come out steaming, bubbling' (similarly, Nez Perce *mú'yn* is derived from *mú'y-* 'to rise as steam,' and Chayahuita *tomontërinso* from *tomoitërin* 'for steam to rise up'). Carrier *yentsel* "steam (over the ground after rain)" contains *yen* 'earth,' Itzaj colexifies 'scab' inter alia, while Santiago Mexquititlan Otomí *h-ñä* is analyzable as 'IMPERSONAL.VOICE-speak.' The term means 'steam' as well as 'voice, language.' Wintu *li'd* also means to "extinguish fire, sprinkle water on hot rocks to make steam, throw water on fire" and *x0s* also means 'lungs' (which is suggestive of an underlying association with 'breath'), Chayahuita colexifies 'steam' with 'airplane,' Guaraní with 'nose, beak, point,' and Hani with 'to cook by steaming.' Bwe Karen *θu* also means 'blood' inter alia, Fijian *kuvu* also denotes "the

foam at the front of a swiftly moving canoe,” and Hawaiian *māhu* also several kinds of trees.

#### 59. *The Straw*

Representation: 44%

Motivated: 70.0%

Thereof Analyzable: 26.2%

Thereof Colexifying: 43.8%

Thereof by Contiguity: 53.3%

Thereof by Similarity: 6.9%

Recurrent associated meanings: grass, dry, maize, chaff, stalk/cane, fodder, pasture/lawn, mat, kindling, hat

By far the most frequent association for ‘straw’ (or ‘hay’) is that with ‘grass, weed,’ either by colexification (sometimes with additional colexified meanings, see section 28), or by morphologically complex terms, which have, in the overwhelming majority of cases, an element meaning ‘dry’ as their second constituent. The association is also common in diachrony in Indo-European (Buck 1949: 520-521). Colexification is found in Buli, Ngambay, Abzakh Adyghe, Badaga, Bezhta, Khalkha, Kildin Saami, Upper Chehalis (by the term *s-máq<sup>m</sup>m=umš* ‘CONTINUATIVE-prairie=place’), Cheyenne, Chickasaw, Highland Chontal, Itzaj, Kashaya, Kiowa, Lakhota, Nez Perce (also colexifying ‘lettuce’ in one of the relevant terms and “bits of straw to start fire” and ‘kindling’ in the other, the latter association is shared with Ancash Quechua), Nuuchahnulth, Oneida, Pipil, Tuscarora, Yaqui, and Mandarin (22 sampled languages; Pipil colexifies ‘straw for hatching’ more specifically, and in Efik, ‘straw’ and a particular type of grass are colexified). Complex terms involving a term meaning ‘dry’ or ‘withered,’ as for instance in Yoruba *koríko gbígbe* ‘grass dry,’ are found also in Nivkh, San Mateo del Mar Huave, Fijian, Great Andamanese, Hawaiian, Samoan, and Tetun (Meyah has *mofombrá efėj* ‘weed dry’). Variants are Kolyma Yukaghir *jaqada:čəd-ulegə* ‘horse-grass’ and Cheyenne *mo’e’évôhkêha’e* /*mo’e’é-hóhkêha’e*/ ‘grass-hat’ (note that in Pawnee, *ka’iilcu* can refer to both ‘straw’ as well as a ‘straw hat’ or ‘straw basket’). One of the relevant Fijian terms *kaunisilāmadū* consists of *kau* ‘wood, stick’ the possessive marker *ni*, *silā* ‘maize-like plant, maize’ and *madu* ‘dry;’ a similar term is also found in Yoruba: *igi agbado tabi bàbà* ‘wood maize or guinea.corn,’ and Carrier has a semianalyzable term where one constituent means ‘wood’ or ‘stick.’ Further, there are also other languages in which ‘straw’ is lexically associated with ‘corn’ or ‘maize’: Wayampi *awasi-i-ε* is analyzable as ‘maize-leaf/stalk-PAST’ (referring to ‘straw of maize’ specifically), Chayahuita has *shi’shi’ sha’huētē pochin ninin-so’* ‘maize leather/bark like do/be-3SG.SUB,’ and Khoekhoe features a semianalyzable term with a constituent meaning ‘corn.’ Greek, Lake Miwok, and Embera colexify ‘straw’ with ‘chaff,’ and Khoekhoe, Muna, and Basque with ‘stalk’ and/or ‘cane’ (‘stalk of corn’ specifically in Khoekhoe). Similarly, Dongolese Nubian has *hasénnkášš(i)* ‘refuse of reaped crop, straw,’ containing *hásed* ‘stumps of crops remaining after harvest, stubble’ and *kášš(i)* ‘refuse, waist (of vegetation).’ In Khalkha and Abipón, presumably by provenience contiguity, ‘straw’ and ‘mat’ are colexified. Aguaruna appears to colexify ‘straw’ with ‘layer of grass’ and ‘lawn,’ Cavineña colexifies ‘pasture,’ and Embera *phōārā*, associated with another gender, also means ‘pasture, grassland.’ In Tetun, the association

is present due to colexification of 'grass' and 'pasture' by an analyzable term with the other constituent meaning 'dry.' Finally, Badaga, Sora, and Guaraní colexify 'straw' and 'fodder.'

Other associations include: the plural form of Buli *wuuk* 'grass, blade of grass, straw' means "bushland, grassland, bush," and Abzakh Adyghe *χə-pq'e* is analyzable as 'mow-place' and also denotes a 'field after harvest.' Blackfoot *soi'stsipínnakssin* is derived from *soi'stsipínnaki* 'to harvest' and indeed also denotes a 'harvest.' Nez Perce *peqes* denotes, alongside 'straw,' the 'Bunchgrass' specifically as well as "wild wheat, wheat." Wappo *pá'ha?* also means 'belt' (presumably due to phonological collapse of Span. *faja* 'strip, waistband' with *paja* 'straw'), and Wintu *sunus* also means 'nest, lair.' Another Wintu term, *tEh*, means inter alia 'bed' alongside 'spread, hay.' Central Yup'ik *cupun* ~ *cup'un* is analyzable as /cupe-n/ 'blow-instrument' and can also refer to 'embers' and a 'rifle.' Miskito *rais mina* is analyzable as 'rice husk' (*mina* in fact also denotes the 'foot' and by extension also the lower part of something), while *waha*, another Miskito term for 'straw,' colexifies 'leaf.' Ancash Quechua *achu* also means 'splinter,' Tsafiki *yaja* 'straw for houses' contains *ya* 'house, roof,' and Bwe Karen *bu ləkhwa* contains *bu* "paddy, unhusked rice;" the language also has another semianalyzable term containing an element meaning "(on) the upper or outer surface of."

#### 60. *The Sun*

Representation in Database: 99%

Motivated: 51.8%

Thereof Analyzable: 5.8%

Thereof Colexifying: 46.0%

Thereof by Contiguity: 25.7%

Thereof by Similarity: 10.9%

Recurrent associated meanings: day, moon, clock, sunshine/sunlight, time, hot/heat, hour, sail, calendar, noon, weather, god, eye, burn, sunray/sunbeam, year, fire

The most frequent association between the 'sun' and other meanings is that with 'day, daytime' by contiguity. Colexification is found in as many as 37 languages, namely Hausa, Mbum, Rendille, Buin, Gurindji, Kyaka, Mali, Ngaanyatjarra, Sko, Toaripi, Japanese, Sora, Kildin Saami, Cahuilla, Ineseño Chumash, Comanche, Itzaj, Kiliwa, Lake Miwok, Pawnee, Quileute, Xicotepec de Juárez Totonac (by the complex term *chi'chini'* /chi'chi-ni'/ 'warm-AGT'), Yana, Yaqui, San Lucas Quiaviní Zapotec, Copainalá Zoque, Lengua, Miskito, Rama, Tehuelche, Wichí, Fijian, Great Andamanese, Hawaiian, Manange, White Hmong, and Sedang. Similarly, Kyaka and Tasmanian (Southeastern and perhaps Middle-Eastern varieties) colexify 'sun' with 'daylight.'

Frequently, 'sun' is also extended to convey other time-related concepts. Indeed, the very notion of 'time (of day)' is expressed by the same term as 'sun' in Ngambay (here also 'moment' is colexified), Buin, Burarra, Badaga, Cahuilla, Itzaj, and Tehuelche, and in Ngambay, Highland Chontal, and Cashinahua (here also colexifying 'year,' as in Wayampi, and 'brilliance, strength of sun'), the relevant terms also mean 'hour' (note that none of these languages also colexify 'day!'). Some languages, namely Burarra, Gurindji, Toaripi,

Upper Chehalis, Cheyenne, Comanche, Kashaya, Kiowa, Wappo, Wintu, Bora, Cubeo, Jarawara, Yanomámi, and Kapingamarangi, employ their word for ‘sun’ directly also to denote a ‘clock’ or ‘watch,’ and in Cheyenne and Wappo also ‘calendar’ (and there are many languages where ‘clock’ is expressed by a complex term involving ‘sun,’ see section 79). In Sko and Buli, ‘noon’ specifically is colexified with ‘sun,’ and Buli and Itzaj feature terms which can also convey the meaning ‘weather’ and/or ‘climate’ (Buli also colexifies ‘season’).

As an alternative to the colexification of ‘sun’ and ‘day,’ there are languages with analyzable terms for ‘sun’ on the basis of ‘day,’ most frequently in Southeast Asia and Oceania (Urban 2010). Sahu has *wangere ma la’o* ‘day POSS eye,’ Yay *taaj’ van* ‘obstruct.view day,’ and Fijian, alongside *sigā*, which colexifies ‘day’ and ‘sun,’ also has the complex term *mata-ni-sigā* ‘eye/face-POSS-day/sun.’ San Mateo del Mar Huave has *teat nüt* ‘father day.’ The association with ‘eye’ is also found in Malagasy (*masonandro*, analyzable as /màso-n-àndro/ ‘eye-GEN-day’), and a somewhat similar term is found in Vietnamese (*mặt trời* ‘face sky’); note also Tetun *loro-matan* ‘sun-eye’ for ‘disk of the sun’ specifically. Furthermore, there is evidence for a diachronic association between ‘day’ and ‘sun’ in Basque (compare *eguzki* ‘sun’ and *egun* ‘day’), and for a connection between ‘eye’ and ‘sun’ in Irish, where the present-day word for ‘sun’ is cognate with the inherited Indo-European word for ‘eye’ (Mallory and Adams 2006: 128). A Middle-Eastern Tasmanian term for ‘sun’ is recorded as *pōganubrenā* which resembles *pōga(na)* ‘man’ and *nūb(ē)rē(na)* ‘eye.’ However, the simplex terms are not attested for the same language or dialect but for different ones, and thus it is unclear whether the mentioned term can be analyzed in this way or whether a diachronic connection between ‘eye’ and ‘sun’ should be recognized. A different recurrent pattern is colexification of ‘sun’ with ‘warm, warmth’ or ‘hot, heat,’ or complex terms exhibiting this association (see Buck 1949: 54 for the possibility of this connection in Irish). Xicotepec de Juárez Totonac *chi’chini* was already mentioned. Similarly, Angkor has *hūfū-hamindi* ‘hot-bone/very,’ the Nunivak island dialect of Central Yup’ik has *puqla-neq* ‘warmth/heat-thing.that.results.from’ (attested presently only in the Nunivak Island dialect, but recorded in the 19<sup>th</sup> century also for other dialects), and Abipón has *m-pae-Ra* ‘POSS.INDEF/3SG-hot-ABSTR.’ This term colexifies ‘sun’ with ‘heat (of the sun),’ and this pattern is also found utilizing monomorphemic terms in Buli, Miskito, Rama, and Hawaiian (among other meanings in this language). Relatedly, Dadibi and Japanese colexify ‘sun’ and ‘fire’ (the Japanese terms are different in prosody though; Dadibi *sia* in addition denotes the sensation of heat from fire or the sun), and in Lavukaleve and Ancash Quechua, the relevant terms have a verbal usage, in which they assume the meaning ‘to burn.’ In addition, Piro has the semianalyzable term *tkatšj* which appears to contain *tšj* ‘fire.’ Finally, many languages of the Americas use the same term for both ‘sun’ and ‘moon’ (and concomitant extensions typical for ‘moon,’ such as that to ‘month’ in some languages). Sometimes languages have complex terms for ‘moon’ based on ‘sun,’ see section 38. This phenomenon is encountered among the languages of the sample in Blackfoot, Carrier, Upper Chehalis, Cheyenne, Chickasaw, Lakhota, Nez Perce, Nuuchahnulth, Quileute, Tuscarora (according to older sources incorporated into the consulted source), Wappo, Wintu, Bora, Cayapa (here colexifying also ‘lowered, landed, fallen’), Cubeo, Hupda, and Macaguán (for further

associations due to the ‘moon’-reading, see section 38). In some of these languages, while the semantics of the colexifying term clearly allows reference to both heavenly bodies, there are complex terms to disambiguate, for instance Lakhota *wí* ‘sun, moon,’ *ápa-wí* ‘day-sun/moon,’ *həhépi-wí* ‘night-sun/moon;’ the complex term for ‘sun’ of this structure in Cubeo can also be used to refer to a diurnal animal.

In ten sampled languages, Buli (also by the variant term *wen-biri* ‘sun-seed’), Dongolese Nubian, Baruya, Ngaanyatjarra, Rotokas, Waris, Basque, Lesser Antillean Creole French, Embera, and Bislama, ‘sun’ and ‘sunshine’ or ‘sunlight’ are colexified, and Dongolese Nubian and Nuuchahnulth colexify ‘sun’ with ‘sunray, sunbeam.’ Finally, Aguaruna *étsa* also is the name of a mythological hero and the god of hunters as well as the name of yellowish bird and fish species (note that Buli and Highland Chontal colexify ‘sun’ with ‘god,’ and that the same term in the former language is also a “religious concept denoting the ‘alter ego’ or ‘personal god’ of an individual”). Similarly, Bororo also uses its term for the ‘sun,’ *meri*, as a name for certain spirits. Hawaiian, Kapingamarangi, and Samoan colexify ‘sun’ with ‘sail’ (according to Elbert and Pukui 1986: 188, in the sense of ‘sun’ the relevant Hawaiian term *lā* goes back to Proto-Polynesian \**la’aa*, and in the sense of ‘sail’ to \**laa*).

Other associations are: Buli *wen* colexifies ‘sun’ with ‘sky’ as well as ‘up, upwards,’ while Koyraboro Senni *woynow* ~ *weynow* ~ *woyna* might contain *woy* ‘woman, female’ (which would then in all likelihood be motivated by mythological complexes; note also *woynow* ‘hemorrhoids’). Buin *rua* also means ‘door’ (accidentally, due to borrowing from English) as well as “Be reddish-yellow (the colour of the sun at dawn),” and Burarra *marnnga* is the only term in the sample which simultaneously denotes both ‘sun’ and ‘star’ (as well as a type of shellfish similar in appearance to a star and several types of sea urchin; the language also has an unrelated monomorphemic term for ‘star’). Kwoma *ya* also means ‘decoration’ and ‘money’ among many other things, Nunggubuyu *alir* is also used to refer to the ‘starfish,’ and Yir Yoront *pung* is said to also mean ‘payback pendant.’ Individual variants of Badaga *ottu* ~ *ottu* ~ *hottu* can also mean ‘to adhere,’ ‘pod, vegetable waste,’ and act as a suffix “indicating numerical frequency.” Greek *ἥλιος* also means ‘sunflower,’ Japanese *taiyō* also ‘ocean,’ Ket *ī* also ‘name,’ Cheyenne *eše’he* also ‘compass,’ and Kiowa *phe* is also used to convey the meaning ‘summer.’ The literal meaning of Nez Perce *wiyetené’t* is ‘the one that travels’ (to have terms for celestial bodies derived from a verb meaning ‘to travel’ is common in the American Northwest). Nuuchahnulth *hupał* also is used to refer to the ‘thimbleberry,’ and Wintu *tulcuheres*, which is at the same time the name of a mythological hero, might be, with additional information from a Wintu myth, analyzable as “the one who was beaten as a (potential) spouse,” compare *tul* ‘to beat a spouse.’ Central Yup’ik (Norton Sound dialect) *macaq* also means ‘to shine.’ Kaingang *rā* also means ‘near close,’ ‘below,’ and ‘to begin, be about to,’ Jarawara *bahi* also denotes ‘thunder’ and ‘lightning,’ whereas the Toba term *i-coṽadelec na* is analyzable as ‘3-illuminate 3sg.’ Wayampi *kwalai* also means ‘dry season,’ Bislama *san* also rarely is used with the meaning ‘son’ (due to collapse of Engl. *sun* and *son*), Hani *naolma* might contain *ma*, a classifier for big things, and Rotuman *asa* also means ‘reputation, honor’ inter alia.

61. *The Swamp*

Representation: 71%

Motivated: 42.0%

Thereof Analyzable: 25.3%

Thereof Colexifying: 16.9%

Thereof by Contiguity: 12.3%

Thereof by Similarity: 3.7%

Recurrent associated meanings: pool/pond/lake, mud, water, puddle, dirt, wet/moist, clay, reservoir, damp, earth, meadow/grass, lagoon, pus

Terms for ‘swamp’ (‘fen,’ ‘marsh,’ ‘bog,’ ‘mire’) are frequently associated with ‘pond, pool,’ or ‘lake,’ either because languages do not distinguish lexically between the two and boundaries are somewhat fluid, or because ‘swamp’ is expressed by a complex term on the basis of one of these meanings. Colexification is found in Buli (where the meaning ‘lake, pond’ is rare), Ngambay, Gurindji (where the relevant term is also the name of a specific lake and denotes “something in ceremony” additionally), Kwoma (also colexifying ‘waterway, canal’ and ‘lagoon,’ which latter meaning is also colexified in Yanomámi, while in Tetun *kolan* ‘lagoon’ is in addition glossed as ‘saltwater swamp’ in parentheses), Basque, Ket, Chickasaw, Lesser Antillean Creole French, Pawnee (by a semianalyzable term containing an element meaning ‘to be a place’), Kaingang, Yanomámi (by the term *wawëwawë*, reduplicated from *wawë* ‘wide, empty’), and Hani (by a semianalyzable term containing an element meaning ‘sea, ocean’ and ‘to soak in water’). Similarly, Sedang colexifies “stagnant pond” more specifically. In the case that analyzable terms are found, the second constituent can mean ‘mud,’ as in Baruya *ara’bunya* /*araka-bunya*/ ‘mud-pond’ and also in Kyaka (there is also a variant term where additional constituents are present), ‘dirt’ as in Hawaiian *ki’o lepo* ‘pool dirt,’ ‘wet’ as in Mandarin *zhao3-ze2* ‘pond-watery.terrain/wet,’ ‘spoiled’ as in Meyah *méren óbóha*, which is apparently analyzable as /*méren obohá*/ ‘lake spoiled,’ or ‘earth’ as in Khalkha *cögerym yazar* ‘pool/lake earth.’ In addition, Ngaanyatjarra *yurungarri* consists of *yuru* ‘lake’ and *-ngarri*, glossed as ‘those associated with,’ and in Ket, there are several terms for ‘swamp’ containing one for ‘lake’ inflected for plurality (see § 4.5.2.1. on the quasi-derivational function of the plural morpheme in Ket). A term for ‘swamp,’ *yoka’ ~ ayoka’*, that is apparently formed by adding *ka* ‘something’ to *yohi ~ ayohi ~ ayohi’ ~ hayo’ha ~ ayo ~ ayox* ‘lake’ is found in Biloxi. There are also several languages in which ‘swamp’ or ‘marsh’ are colexified with ‘puddle,’ namely Muna, Basque, Khalkha (“mud puddle” more specifically), San Mateo del Mar Huave (the relevant term *ndorrop yow* is analyzable as ‘hole water’), the Cuisnahuat dialect of Pipil (by the analyzable term *ta:l-a:pu:ni* ‘ground-water-be.born,’ which also means ‘spring of water,’ compare Upper Chehalis *mólx<sup>w</sup>* ‘swamp’ and *mó’l-* ‘spring’), Arabela, and Bora. A Chickasaw term for ‘swamp’ has the idiolectal meaning ‘puddle.’ Further, in Miskito, *piahka tara* is analyzable as ‘puddle big,’ and in Hawaiian *ki’o lepo* as ‘pool/puddle dirt.’

Several of the associations arising in complex terms with ‘lake’ also occur in other configurations. Buli, Hausa, Yoruba, Kosarek Yale, Basque, Khalkha, Arabela, Hupda, Ancash Quechua, Hawaiian, and Lenakel directly colexify ‘swamp’ with ‘mud’ (“[m]ud which is so deep as to be practically impassable” in Hausa and ‘black mud’ specifically in Hupda), Baruya has *ara’darya* /*araka-darya*/ ‘mud-area,’ Kyaka *ipwua manduwua* ‘water



mud' (this term also denotes 'watery mud' and 'slush'), and San Mateo del Mar *n-ajndor iüt* /*na-ajndor iüt*/ 'AGT-be.muddy-ground.' Further, Cubeo *yao-bu* is derived from *yaa* 'mud, adobe' by means of the classifier *-bu* for hard and/or round objects, Kaingang has *óré ki goj* 'mud in water/river,' and Fijian *vanua oruoru* 'land muddy' (Fijian also has *lōlōbo*, reduplicated from *lōbo* 'to stick in the mud, to penetrate something'). Furthermore, Guaraní *tuju rapo* and *tuju rupa* contain *tuju* 'mud, abode.' The relevant Khalkha term also colexifies 'clay,' so do the Buli and Ancash Quechua terms. Likewise, terms for 'swamp' in which the meaning 'earth' figures are not only attested in conjunction with 'lake': Efik has *memmem isōñ* 'soft earth/ground,' and in Embera, *egoró-susúa* contains *egoró* 'ground, earth' and *susúa* 'watering place,' a semianalyzable term including a constituent with the meaning 'earth' is also found in Bora. Alongside the association with 'dirt' by the overt term in Hawaiian, Buli, Kosarek Yale, and Rama colexify 'swamp' with 'dirt.' There are also many languages with complex terms for 'swamp' where one of the constituents means 'water,' some of which were already mentioned above. Kanuri has *nji-bòné* 'water-lie.down' and Kyaka *ipwua manduwua* 'water mud,' *ipwua wara pete* 'water watery.mud pool' and *ipwua/ipya inju malu singi dokona* 'water mud much usual.location there.' Ngaanyatjarra *kapingarri* (meaning "area where water runs, swampy area") consists of *kapi* 'water' and *-ngarri* 'those associated with,' Abzakh Adyghe *psə-č'e=psə-lə* is analyzable as 'water-source=water-DERIV,' Ket *ulteyin* as /ul-te-in/ 'water-lake-PL,' Sora *dunḍam'da:lo:n* as /dunḍam-d'a:-'lo:-n/ 'get.out.of-REFLX-water-ground-N.SFX,' Haida *xawtl'adaanggaa* 'be swampy' contains *xaw* 'liquid' and *daanggaa* meaning 'have been discarded' inter alia, Kashaya *q'at'u?ul* is analyzable as /ahq<sup>h</sup>a-t<sup>h</sup>u?ul/ 'water-old' (and also denotes "stagnant water" generally), Nez Perce *kus'ín wé'tes* as /ku's-í'n wé'tes/ 'water/dew-with land,' Pipil (Cuisnahuat dialect) *tal-a:pu:ni* as 'ground-water-be.born,' Wintu *me'm lat-i* as 'water/river wet/damp/moist/soak/drench-??,' Yaqui *ba'a jeelo* as 'water near,' and Wichí *inot-w'et* as 'water-place,' semianalyzable terms where one of the constituents can be identified to refer to 'water' are found in Blackfoot, Chickasaw, Comanche, Kashaya, Yuki, Cavineña, and Piro.

Furthermore, 'swamp' is associated with 'meadow' and/or 'grass' by colexification in Badaga and Nez Perce (in Badaga also with 'water channel' and "upper part of a ridge;" somewhat similar is White Hmong *hav iav* 'valley grassy' meaning 'wet lowland'). Lake Miwok colexifies 'to be swampy' with 'to be damp,' and in two sampled languages, Chukchi and Wintu, there are analyzable terms for 'swamp' in which the meaning 'damp' figures: *ilā-lqen* 'damp-on.top' and *mem lat-i* 'water/river wet/damp/moist/soak/drench-??' respectively (compare also Japanese *shitchi*, analyzable as /shitsu-chi/ 'moist-place,' Mandarin *zhao3-ze2* 'pond-watery.terrain/wet,' and colexification of 'swampy' and 'wet' inter alia in Hawaiian). In two sampled languages, Central Yup'ik and Rama, a metaphorical association with 'pus' is found: in Central Yup'ik by the analyzable term *imarrluk*, consisting of *imaq* 'contents, bullet, pus, ocean' and the postbase (see § 4.4.2.) *-rrluk* 'one that has departed from its natural state,' and in Rama by colexification (also with 'dirt'). Finally, Badaga, Basque, and Khalkha, colexify 'swamp' with 'reservoir.'

Other associations include: Buli *biung* is primarily the name for a "watering hole that dries up quickly," and *viak* also means 'valley.' Hausa *damba* is also used with the

meaning “a trap in speech” inter alia, Ngambay *bbar* also means ‘sound, noise,’ and Swahili *kinamasi* also ‘slime.’ Berik colexifies ‘swamp’ with ‘delta,’ and Buin *kapunuka* is also an “insulting term for a woman.” Burarra *mugurmulmul* is analyzable as ‘CLASS.DOMESTIC-paperbark.tree’ and denotes a “swamp area” and “in a stand of small paperbark trees,” and the Kwoma term *biira-biira she* is analyzable as ‘open-RED faeces/ash.’ Yir Yoront *walq* also means “hollow place, concave surface” and “main (deep) waterhole.” Badaga *ole ~ hole* also means ‘river’ and Basque *istil* also ‘waterhole.’ Khalkha *sibar namuy* also denotes ‘plaster, stucco,’ and *taca* also ‘broom’ and the ‘Golden Chain’ (*Cytisus laburnum*). Chickasaw *okpachalhlhi* also means ‘mudhole.’ Welsh *siglen* is derived from *sigl* “a shaking, oscillation.” Itzaj *k'al* also means “closing, closed in place,” and the Tuscarora root *-hnaw-* occurring in *uhnàweh* ‘morass, swamp’ also yields the meaning ‘current of water,’ in particular when occurring incorporated. Central Yup'ik (Yukon dialect) *puglerneq* is analyzable as /puge-ler-neq/ ‘come.to.surface-suddenly-thing.that.results,’ and for the Bristol Bay dialect term *angayaq*, compare *angala-* ‘to flutter, wobble.’ Tsafiki *mudú* appears to be analyzable as /mu-du/ ‘achiote-mountain,’ and Tehuelche colexifies ‘swamp’ with ‘skullcap’ or ‘wit’ (original gloss is ‘mollera’). Hawaiian *pohō* also means ‘sunken, sinking’ inter alia, *naele* also includes ‘rock, crevice’ in its denotational range, and *nenelu* also means ‘flabby fat’ or ‘soft,’ again among other meanings. Malagasy *hòraka* also denotes a “wet rice field,” and Manange *Its'ho* also a ‘rope.’ Samoan *taufusi* is also used to refer to a “patch of ground irrigated for the purpose of growing taro,” and *pala* also means ‘(be) rotten’ and ‘to decay, perish’ inter alia when used as a verb. Sedang *lông* also means ‘to sing a lullaby, put a child to sleep, coax.’

## 62. The Tail

Representation: 95%

Motivated: 21.5%

Thereof Analyzable: 3.8%                      Thereof Colexifying: 17.7%

Thereof by Contiguity: 5.6%                      Thereof by Similarity: 14.9%

Recurrent associated meanings: end, penis, tailbone, buttocks, back, train of dress, tail of cloth, tail of coin, tailfeathers

While there are some morphologically complex terms for ‘tail’ in the languages of the sample, recurrent associations are almost exclusively realized by colexification. Most commonly, languages use their word for ‘tail’ also for ‘end’ in general (paralleled in the diachrony of Indo-European in Irish, Buck 1949: 209). This is found in twelve sampled languages, namely Buli (where the relevant term also is the name for the ‘fly-whisk’ inter alia), Yoruba (also colexifying ‘completion, conclusion’), Kwoma (colexifying “lower end” or “downhill side,” as well as “the lower of two entities” more specifically), Kosarek Yale, Abzakh Adyghe, Basque, Khalkha, Lesser Antillean Creole French, Maxakalí, Rama, Kapingamarangi, and Samoan, in which latter the relevant term also means ‘to finish’ as a verb inter alia. Similarly, Sedang colexifies ‘tail’ with “end of cord, loincloth,” and Efik *ntañ ètak* is analyzable as ‘feather end.of.body’ (this term denotes the tail of birds specifically, for tails of quadrupeds and fish, there is a separate monomorphemic term. It should

be noted that in a number of sampled languages, different unrelated terms are used for the tails of specific types of animals). Moreover, Samoan *si'usi'u* is reduplicated, the base being *si'u* 'tip, extremity.' More specific associations go in the same direction: Rendille, Basque, and Kolyma Yukaghir colexify 'tail' with 'buttocks' (furthermore, Ngaanyatjarra colexifies 'tail of insects' specifically with 'bottom, buttocks,' compare also Pipil (Cuisnahuat dialect) *-kwitapil*, which is derived from *-kwita* 'excrement' by means of the (frozen) diminutive suffix *-pil*). Moreover, in Khoekhoe, the same root yields the meanings 'tail' and 'buttock,' with different nominal designants distinguishing the two. In Rendille, Abzakh Adyghe, and Wichí, terms for 'tail' also may refer to the 'back' or 'backside' of something, and Greek and Hawaiian use the same term for 'tail' and 'train of dress.' Similarly, Hausa and Abzakh Adyghe colexify 'tail' with 'tail of cloth' (as well as "a small quantity of sour milk given free to a purchaser of butter" in the former language). The relevant Hausa term is also extended to 'penis,' a pattern also attested in Koyraboro Senni, Lavukaleve, and Cashinahua, and perhaps in Angkor, where this meaning is marked with a question mark in the consulted source (compare the cognacy of New Persian *dum* 'tail' with Old High German *zumpfo* 'penis' noted by Buck 1949: 210; this pattern seems to be widespread in languages of Europe generally, for instance, Latin *penis* originally had reference to the tail of animals). There also is a term for 'tail' where the identifiable constituent is that for 'penis' in Berik. Rotuman *reu* also denotes the 'tail of a coin' (and also the 'foot of a bed'), and a term for 'tail' in Hausa is similarly used to refer to specific motifs on the back side of coins. Relevant terms in Cheyenne, Haida, Macaguán, and Yanomámi simultaneously denote 'tail' and 'tailbone, coccyx' (and the Yanomámi term also 'stinger'), while Nez Perce and Wayampi colexify 'tail' with 'tailfeathers.'

Other unsystematic associations include: Buli *jiuk* also denotes a "bird's trap made of grass," Dongolese Nubian colexifies 'tail' with 'to wash' and 'to send,' while Yoruba *apa èhin ohunkóhun* is analyzable as 'part/arm back whatever.' Baruya *suya* also means 'urine,' and for Kaluli *waf* 'tail,' compare *wafe* 'worm.' As a verb, Muna *punda* also means 'to jump,' and *lensi* also 'to untie, undo, loosen,' while relevant Tasmanian terms in all varieties except the Northern one appear to denote any excrescence of the body: 'wart,' 'scar,' 'wrinkle,' and 'tail.' Sko *pú* is also the name of a furry mammal, and Yir Yoront colexifies 'tail' with "tail-end of spear." The Abzakh Adyghe term *q<sup>ə</sup>eps* is analyzable as /q<sup>ə</sup>(e)-ps(e)/ 'branch/pointed.object-string' and colexifies 'strap' and 'handle,' and another term in the same language can also refer to an 'egg,' a 'seed,' and other things. Badaga *ba:lu* also is used with reference to a 'meteor,' "slender means," and "anything meager." Basque *buztan*, colexifying 'end' and 'butt,' also means 'shoot, sprout,' and *isats* also 'broom' and 'consequence, upshot.' Chukchi *nojɲən* also means 'rump,' whereas Greek *ourá* also means 'queue.' Welsh *llosgwrn* appears to contain *llosg* 'burning,' Cheyenne *hehévá'xe* is also the name for a piece of meat, the 'oxtail.' San Mateo del Mar Huave *wiül* also means 'vixen,' and *coy* also denotes the 'coney' as well as a picture or drawing of any animal and 'rheumatism.' The Nuuchahnulth term *nača* also means 'fluke,' Santiago Mexquititlan Otomí *ts'u* also 'nipple, teat,' and Central Yup'ik (Norton Sound dialect) *pamyuk* ~ *pamsuk* also 'tail of canoe' and 'chorus of song.' Carib *-antikili* is analyzable as /-anti-(e)kili/ 'behind-spin' (compare the origin of German *schwanz* from *sweifen* 'turn around,' Buck 1949: 210). Guara-

*ní tuguái* may also refer to an ‘appendix’ and metaphorically to ‘family, offspring,’ Ancash Quechua *chupa* also is used to refer to a “person who is always behind someone else,” and Tehuelche *t'er* ~ *t'er* also means ‘bristle.’ Fijian *bui* is also used for ‘grandmother’ and ‘old gossip,’ Hawaiian *huelo* also has the figurative meaning ‘inferior,’ and Bwe Karen *ká=mè* is analyzable as ‘hind.part=knot/joint.’ Lenakel *nipik*- also denotes the ‘tail of a stingray’ and ‘stern of canoe or boat,’ and Manange *Imē* also means ‘fin.’ Sedang *tíng* also means ‘to sacrifice to the spirits,’ *xô'i* has similar other meanings, Tetun *ikun* is also used as a term for the youngest of somebody’s children, Vietnamese *đuổi* also means ‘to pursue,’ and Yay colexifies ‘tail’ with ‘head of rice.’

### 63. *The Thorn*

Representation: 72%

Motivated: 42.2%

Thereof Analyzable: 7.8%

Thereof Colexifying: 34.4%

Thereof by Contiguity: 8.0%

Thereof by Similarity: 18.4%

Recurrent associated meanings: needle/awl, splinter, quill, point, sharp, stinger, thorny plant, sticker, tooth, barb, beak, injection, penetrate/pierce, nail

Most frequent cross-linguistically is colexification of ‘needle’ or ‘awl’ and ‘thorn’ (or ‘spine,’ ‘prickle’), which may be either due to provenience contiguity or to perceptual similarity. This pattern is attested in Anggor, Baruya (in these languages colexifying also ‘injection,’ ‘give an injection, inject’ is also colexified in Samoan), Buin, Khalkha, Abipón, Bora, Chayahuita, and Yanómami (in the latter language, the relevant term also means ‘pin,’ Tehuelche *xolnwe* ~ *xoln*, furthermore, contains *xol* ~ *xo:l* ~ *?exol* ‘to sew’ and colexifies ‘nail’ in addition, as does Bislama).

Otherwise, colexification with other sharp pointed objects is frequent. Kyaka, Muna, Nunggubuyu, Kolyma Yukaghir, Cheyenne, Wintu, and Tetun colexify ‘thorn’ with ‘splinter’ (Kyaka also with ‘excrescence’ generally), Kaluli, Kosarek Yale, Highland Chontal, Bororo, and Samoan with ‘stinger,’ and Kyaka, Kosarek Yale, and Lengua with ‘tooth’ (Kyaka also with “biting, erosive” and ‘food’ inter alia). Note also that Maxakalí *xāpxox* presumably contains *xap* ‘stone, bead, seed’ and *xox* ‘tooth, sharp fragment.’ Hausa, Ngambay, Haida, Lesser Antillean Creole French, Central Yup’ik, and Yanomámi colexify ‘thorn’ with ‘quill,’ Kyaka, Bororo, and Lengua with ‘beak,’ Cahuilla, Lake Miwok, Pawnee, and Wintu with ‘sticker,’ while the relevant terms in Kosarek Yale, Bororo, Guaraní, and Miskito are extended to refer to the ‘point’ of an object more generally, and Ancash Quechua also to ‘anything pointed.’ In addition, Yoruba colexifies ‘thorn’ with ‘bone,’ while Cayapa colexifies ‘thorn’ with ‘slim and acute bone’ more specifically.

Efik *ñku'kīm* (containing *kīm* ‘to pierce’) is also used to refer to a ‘spine on a shell’ (alongside ‘patchwork’), Basque *arantza* also means ‘spine of an animal,’ and Hausa *k'aya* also has the meaning ‘fishbone.’ Koyraboro Senni *karji*, Hawaiian *kukū* as well as Samoan *tala* also mean ‘barb’ (*kukū* also means ‘burr’ as well as ‘to hurt by a thorn’ and ‘to hit’ inter alia, and *tala* also ‘prong’ and ‘spur’), and Ngambay’s *hay* also is the term for a ‘particular kind of straw’ (alongside ‘paddle’). Ngaanyatjarra colexifies ‘spike’ (compare Samoan *tuitui*

‘thorn’ with *tui* ‘to stab, jab,’ ‘spike’). Kyaka and Guaraní colexify ‘horn,’ Rotuman *kō* colexifies ‘fork’ (a pattern shared with Samoan, *kō* also means ‘to stab, pierce,’ and the relevant Samoan term also has other meanings), and Bislama *nil* colexifies ‘cock’s spur.’ For Lake Miwok *kīli*, which means “to hook with the horns” as a verb, compare *kīli* ‘horn, antler;’ the Toba term *le* also means ‘tooth of a comb’ and ‘point of a lance.’ The semantic feature of ‘sharpness’ which in all likelihood underlies these patterns of colexification is made explicit in other languages by complex terms. Upper Chehalis *tač=áxn* is analyzable as ‘sharp=edge’ (and indeed may also refer to a ‘sharp edge’), Chickasaw *naa-haloppa-* as ‘something-be.sharp-NMLZ’ (note also that there is a semianalyzable term involving a constituent meaning ‘thing’ in Bwe Karen), and Hawaiian ‘oi’oi is republished from ‘oi ‘sharp’ (this term may also be used to refer to a ‘superior person’ and assumes the meaning “to protrude, stick or jut out” in verbal usage). A semianalyzable term containing a lexical element with the meaning ‘sharp’ is also found in Pawnee, and Kyaka and One directly colexify these meanings. Relatedly, Rotokas and Xicotepec de Juárez have terms for ‘thorn’ derived from verbs meaning ‘to penetrate’ and ‘to pierce’ respectively; direct colexification of these meanings is found in Rotuman, and a semianalyzable term of this type is found in Efik. Another recurrent pattern is colexification of ‘thorn’ with a thorny plant, either with a particular one or generically. Thus Buli *mung* also denotes a thorn tree of the *Acacia* genus (alongside a kind of beetle), Basque colexifies ‘thorn’ with ‘hawthorn,’ for Kiowa *sęigā’t* compare *sęigā* ‘cactus, peyote,’ Tuscarora and Miskito colexify ‘thorn’ with ‘thistle,’ Tehuelche *čorč* ~ *čorč*’ ~ *č’orč*’ ~ *čorč* also denotes the ‘Calafate shrub’ (*Berberis microphylla*), and Pawnee, Wintu, Bora, and Wayampi colexify ‘thorn’ with ‘thorny plant, thorny bush’ in general.

Other associations include: Muna *kiri*, as a verb, also means “to insert a thorn into something” as well as ‘to scrape off,’ and Ngaanyatjarra *ngunyarrma* also means ‘rasp.’ Kosarek Yale *si* also means ‘name,’ and *alok*, another Kosarek Yale term, also ‘for earth to slide off’ and ‘make hollow underneath something.’ Kolya Yukaghir *nono* also means ‘handle,’ Cahuilla *čujal* also ‘jumping cactus,’ and Bororo *oto* also ‘peak, front.’ Cavineña *acui-ja* is analyzable as ‘tree-GEN,’ and Miskito colexifies ‘thorn’ with ‘twinge.’ A literal translation of the Rama term *kú up* (/ku up/) would be ‘bird’s.wing eye,’ and Tsafiki *po* is also used for trees of the *Guadua* genus. Fijian *voto* also means ‘root of a body hair’ and ‘a hundred voivoi leaves.’ Malagasy *tsilo* also denotes the ‘needles’ of the pine tree, and Tetun *aitarak* also means “prickly, rough.”

#### 64. The Thunder

Representation: 91%

Motivated: 41.5%

Thereof Analyzable: 17.9%

Thereof Colexifying: 24.5%

Thereof by Contiguity: 12%

Thereof by Similarity: 28.1%

Recurrent associated meanings: lightning, god/spirit, sky, roar, noise/sound,  
(thunder)bird, cry/wail, storm, cloud, gun, rain, electricity, voice

21 sampled languages, Buli, Ngambay, Rendille, Yoruba, Kwoma, Abzakh Adyghe, Japanese, Itzaj, Xicotepec de Juárez Totonac, Yana, Yaqui, Central Yup'ik, Abipón (by the analyzable term *kahag-Ran-Ra* 'to.lighten-CAUS-ABSTR'), Arabela, Aymara, Hupda, Jarawara, Tehuelche, Bislama, Bwe Karen, Takia, and perhaps Middle-Eastern Tasmanian, directly colexify 'thunder' with 'lightning, thunderbolt' (compare the semantic shift from the latter to the former in Lithuanian, Buck 1949: 58). However, in contrast to the association with 'lightning,' complex terms for 'thunder' on the basis of 'lightning' are much rarer. San Mateo del Mar Huave *ajüy teat monteoc* is analyzable as 'walk father thunderbolt,' and Lavukaleve and Rama have semianalyzable terms for 'thunder' containing the respective term for 'lightning.'

For 'thunder' specifically, a general recurrent pattern is seen in terms derived from terms denoting some kind of loud noise (see Buck 1949: 57 for details on similar evidence from Indo-European languages). Terms in many languages contain a verb meaning 'to roar.' Mbum has *balà-mbàm* 'roar-rain' and *fómà-mbàm* 'scolding-rain,' Guaraní has *arakororō* 'sky-growling,' and Manange *3mo putul njut-si*, involving the constituents *3mo* 'sky,' *putul* 'dragon,' and *njut* 'roar.' Khoekhoe, Buin, Nunggubuyu, and Sora directly colexify 'thunder' with 'roar,' 'growl' and/or 'boom' (Sora also 'to shake'), and a semianalyzable term containing an element with that meaning is also found in Great Andamanese. Similar terms, in which, rather than 'growl,' more general terms for 'noise' or 'sound' figure as constituents are Katcha (*thimpido*) *kafara* '(sky/rain) make.noise/cry,' Meyah *mocgój ogúgur* 'cloud noise,' Japanese *kami-nar-i* 'god-sound-NR,' and Tetun *rai-tarutu* 'earth-noise.' In Santiago Mexquititlan Otomí, Copainalá Zoque, and Bororo, the meanings are colexified, and semianalyzable terms with a constituent meaning 'noise' are featured in Copainalá Zoque and Abipón. Another class of motivated terms relating to some sort of noise is that comparing the sound of 'thunder' to 'wailing' or 'crying.' Katcha has *thafara ma thimpido* 'wailing GEN sky/rain,' and Comanche *tomo-yaketu* contains *tomo* 'cloud, sky' and *yaketu* 'cry.' The Kiliwa term *?+kwi-y=h+mii* has the same structure: it is analyzable as 'DN+cloud=3+cry,' White Hmong has *xob quaj* 'Xob cry' and *xob nroo* 'Xob moan,' and the meanings are colexified in Yir Yoront. In Carrier, the word for 'thunder,' *t̥t̥ni tetni*, contains *t̥t̥ni*, the name for a "gigantic bird" in Carrier mythology and *tetni* 'cry' (it is likely that *t̥t̥ni* is the Carrier incarnation of the 'thunderbird'). Similarly, Lakota has *wakíyqhothúpi*, literally 'thunderbirds call.' Colexification of 'thunder' with a (mythological thunder-)bird is also attested in Waris, Upper Chehalis, Haida, and Miskito. Kiliwa also has an alternative term: *ha?=kw-?-ny+mar=kwi-y* 'voice=WH-DN-POSS+image=cloud.' Similarly, in Pawnee the 'thunder' is called *wakuhtakaahak*; this term is analyzable as /wakur-tahaahak/ 'voice-drop.down. Moreover, Bakueri *ngálá lówa* is analyzable 'gun god/sky' and, paralleling this association, Lavukaleve and Rotokas terms colexify 'gun' and 'thunder' (compare the extension to 'cannon' in Romanian, Buck 1949: 58). In Efik *erituak'* *en'yōn* is analyzable as /erituak' en'yōñ/ 'knocking/beating.of.drum sky.' Still further, Hausa *aradu* also dialectally denotes "the wedge used in splitting palm wood," and *tsawa* also is used to refer to a 'loud rebuke' or "the cracking of newly-burned pots for no apparent reason" inter alia, while *cida* is also the name of a spirit, again inter alia. In Buin, *kururu* is also used for a "rumbling noise" in general and is also the name for a "large wooden trumpet" and a

“large buzzing fly,” Bezhta has *hasa qäļeyoti* ‘sky:ERG shout:MASDAR,’ and the Basque term *dunbots* also may refer to a ‘din’ or ‘clamor.’ Cheyenne *ma’heo’o énéstoohé* contains elements meaning ‘god’ and ‘call.’ Ineseño Chumash *’ašaqšk’apš* revolves around the verb *šaqšk’ap* ‘to clap.’ The final element is unknown. The Wichí term *pelhay y’iplhi* contains the lexical elements *pelhay* ‘storm’ and *y’ip* ‘sing’ (note that the same term in Embera yields the meanings ‘thunder’ and ‘storm, tempest’ when associated with different genders, and that ‘storm,’ ‘rainstorm,’ or ‘thunderstorm’ is colexified with ‘thunder’ in Ngambay, Rotokas, and Carib). A figurative Hawaiian term for ‘thunder’ is *’u’ina pōhaku a Kāne*, literally ‘crackling rocks of Kane’ (Kane being the principal Hawaiian god), with the element *pōhaku* also meaning ‘thunder’ by itself. Underlying this denomination, as well as *pohā-ka’a* ‘crack-turn’ (though note for the first element that *pōhaku* has a short form *pōhā*) is the belief that thunder was caused by the gods rolling around rocks in the sky. As the discussion so far shows, in complex terms, ‘cloud,’ ‘sky’ or ‘rain,’ as meanings contiguous to ‘thunder,’ frequently figure in complex terms for that meaning. Alongside the terms in Meyah, Comanche, and Kiliwa, One, too, has a term involving cloud: *yemi piyale* ‘cloud break’ (and note that Badaga *iḍi* can also mean ‘to break’ inter alia). Alongside Mbum and Katcha, ‘rain’ also figures in Yaqui *yuku jimaa-ri* ‘rain throw-RES,’ Buli *ngmoruk* ‘rain’ rarely assumes the meaning ‘lightning, thunder,’ and Hawaiian colexifies ‘thunder’ with “raindrops, fine rain, to rain gently” inter alia, while Noni has a semianalyzable term involving ‘rain,’ and Aguaruna *ipamát* also denotes “to rain as a sign of death or an imminent attack” as well as ‘to reveal a message in dreams.’ Alongside Efik, Katcha, Bezhta, Guaraní, and Manange, ‘sky’ is also the meaning of one of the constituents of Hani *aoq-jiq* ‘sky-sift.’ An association with (a) god or a spirit is, next to Bakueri, and Cheyenne (where ‘battery’ is in addition colexified), also found in Japanese (*ika-zu-chi* ‘horrible-GEN-spirit,’ this term is archaic and the accuracy of the morphological analysis is questionable) and Yay (*pya<sup>3</sup> ray<sup>4</sup>* ‘spirit cry’). Further, in Nez Perce, where *hinmé’t* is also the name of the “spirit of a cloud that makes noise” (alongside the reading as ‘electric storm’) as well as in Yana, Biloxi, Cheyenne, Tuscarora, Wayampi, and Yanomámi, there are associations with some kind of god or spirit by colexification. In Chayahuita, the word for ‘thunder’ is also the name of the person who castigates the evil after their death.

Central Yup’ik and Hupda colexify ‘thunder’ with ‘electricity,’ which is in all likelihood a sideeffect of the fact that they also colexify ‘lightning’ using the same term; the association is much more frequent for ‘lightning,’ compare section 35.

Other associations include: for Angkor *burihoai*, compare *buri* ‘many things existing upright.’ *Nimamindohoafi* ~ *nimamindihoafi* contains *hoafi* ‘to talk;’ a literal translation provided by lexicographers is “above possessive taik [sic!].” Kyaka *yungala* also means ‘praying mantis,’ and Muna *tondu* also ‘to sink, drown.’ Rao *gramvuvre* appears to contain *gra* ‘sun,’ and Rotokas *varake-oto*, meaning ‘thunder’ as well as ‘thunderstorm,’ seems to be analyzable as ‘very.high-to.punch.’ Sentani *ku* also means ‘bracelet,’ and Blackfoot *kšiistsikomm* contains *kšiistsikó* ‘day.’ For Pawnee *kirir*, compare *kirir* (*uur...*) ‘to shake, tremble.’ Wintu *t<sup>h</sup>um* is also used with the meaning ‘coo.’ Copainalá Zoque colexifies ‘light thunder’ with ‘spark,’ and Jarawara *bahi* also denotes the ‘sun’ (and, departing from there, ‘clock, watch,’ see Dixon 2004: 71), whereas Rama *dama yatangi* contains *dama* ‘grandfa-

ther.’ Fijian *kuru* also means ‘to pursue’ and ‘to jostle,’ Hawaiian *hekili* is metaphorically extended to also mean ‘passion, rage’ and also denotes a yam species. Samoan *fāititili* contains *tili* ‘to tremble.’ *Ta’a-lili* is fully analyzable: ‘be.loose-tremble.’

#### 65. *The Tree*

Representation: 97%

Motivated: 70.3%

Thereof Analyzable: 5.4%

Thereof Colexifying: 65.3%

Thereof by Contiguity: 50.8%

Thereof by Similarity: 0.9%

Recurrent associated meanings: wood, stick, plant, trunk/log/pole, shrub/bush, forest, branch, wooden artifact, thing, bone, splinter, canoe, gun/rifle

By far the most frequent lexico-semantic association is that with ‘wood’ (see also Buck 1949: 48 for this pattern in Indo-European). It is realized most frequently by colexification, in 83 languages of the sample, that is, in a little less than sixty per cent of sampled languages. This figure draws close to that arrived at in an earlier survey by Witkowski et al. (1981), who report that the pattern is found in two thirds of their sampled languages. In the present sample, colexification of ‘tree’ and ‘wood’ is found in Efik, Hausa, Khoekhoe, Koyraboro Senni, Ngambay, Swahili, Yoruba (also colexifying ‘fuel’), Berik, Buin, Burarra, Dadibi, Gurindji, Kwoma (where the relevant term may also refer to a “wooden beater,” “wood-carving,” and “slit-drum, hollow log drum”), Kyaka (the term also means “human, earthly, mortal” and ‘below, lower’), Mali, Ngaanyatjarra, Nunggubuyu, Meyah, Tasmanian (all varieties except the Northern one, for which data are lacking), Toaripi, Sahu, Sko, Waris, Kosarek Yale, Badaga, Japanese, Ket, Khalkha, Nivkh (colexifying ‘firewood’ and ‘wooden’ more specifically), Kildin Saami, Welsh, Kolyma Yukaghir, Biloxi, Cahuilla, Carrier, Upper Chehalis, Chickasaw, Highland Chontal, Ineseño Chumash, Comanche, San Mateo del Mar Huave, Itzaj, Kiowa, Lakota, Nez Perce, Santiago Mexquititlan Otomí, Pawnee, Pipil, Xicotepec de Juárez Totonac, Wappo, Yana, San Lucas Quiaviní Zapotec, Copainalá Zoque, Aguaruna, Bororo (by the term *iguru*, for which compare *igu* ‘rope’ and *ru* ‘fire’), Cashinahua, Cavineña, Cayapa, Embera (where the meanings are associated with different genders), Guaraní (where the relevant term can also refer to a ‘plank of wood’), Huambisa, Hupda, Jarawara, Kaingang, Macagúan, Miskito, Piro, Tsafiki, Wayampi, Wichí, Yanomámi, Fijian, Hawaiian, Bwe Karen, Lenakel, Malagasy, Mandarin, White Hmong, Rotuman, Samoan, Takia, Tetun, Yay, and Bislama. In addition, there are a few sampled languages in which ‘tree’ is expressed by a morphologically complex term involving a constituent meaning ‘wood.’ Lesser Antillean Creole French has *pié-bwa* ‘stem/foot-wood’ and Manange *ʔŋ-3tuŋ* ‘wood-copse/trunk.’ Witkowski et al. (1981) propose that complex terms for ‘tree’ involving ‘wood’ can be seen as a quasi-evolutionary development, in which ‘wood’ is the more “salient” referent in small-scale societies which is first extended to ‘tree’ by colexification and then the ‘tree’-reading is singled out by complex terms as societal complexity increases. In the light of this hypothesis, it is interesting to note that a Creole language in the present sample features such a complex term, which would entail that, if their general scenario is correct and also applicable to Lesser Antillean Creole



French, the development must have occurred in a very short time span, given that the lexifier language French does not colexify 'tree' and 'wood' (another possibility would be that Lesser Antillean Creole French acquired this pattern through relexification). In addition, in Upper Chehalis, *łíšałíš* 'a clump of trees, woods' is reduplicated from the root *łíš* 'wood,' in Cahuilla, *kélawat* 'tree, wood' is derived from the verb *-kélaw-* 'to gather wood,' and in Cubeo, 'tree' is *jocw-cw*, consisting of *jocw* 'wood' and the classifier for tree-like objects *-cw*.

Colexification of 'tree' with parts of trees also occurs. Mirroring the complex terms in Lesser Antillean Creole French and Manange, 'tree' and 'trunk,' 'log,' or 'pole' are colexified in 19 languages, namely Mbum, Kwoma, Ngaanyatjarra, Badaga, Chukchi, Ket, San Mateo del Mar Huave, Nivkh, Kiowa, Nez Perce, Oneida, Xicotepec de Juárez Totonac, San Lucas Quiaviní Zapotec, Aguaruna, Jarawara, Yanomámi, Hawaiian, and Kapingamarangi, 'tree' and 'stick' are in 34 languages (Efik, Khoekhoe, Ngambay, Noni, Burarra, Gurindji, Kwoma, Lavukaleve, Ngaanyatjarra, where "magic stick" is also colexified, Nunggubuyu, Yir Yoront, Chukchi, Ket, Khalkha, Kolyma Yukaghir, Biloxi, Blackfoot, Carrier, Upper Chehalis, Chickasaw, Comanche, Kiowa, Wappo, Yana, San Lucas Quiaviní Zapotec, Guaraní, Jarawara, Lengua, Maxakalí, Rama, Fijian, Hawaiian, Kapingamarangi, and Sedang), and 'tree' and 'branch' in Ngambay, Buin (colexifying 'small tree' and a particular tree species more precisely), Welsh, Blackfoot, and Kiliwa. Ngaanyatjarra and Hawaiian colexify 'tree' with 'splinter,' and Abipón is unique in colexifying 'tree' and 'bark' (as well as 'rose bush').

In Khoekhoe, Ngambay, Rendille, Ngaanyatjarra, Yir Yoront, Sora, San Mateo del Mar Huave, Kashaya, Kiowa, Nuuchahnulth, Quileute, Yaqui, Ancash and Imbabura Quechua, Fijian, Hawaiian, Rotuman, Samoan, Takia, Tetun, Yay (here the relevant term also means 'older brother'), and Bislama, 'tree'-terms are also used for 'plant' in general (though the Rendille term explicitly excludes 'grass' from its denotational range); similarly, Koyraboro Senni, Rendille, Kyaka, Nunggubuyu, Toaripi, Badaga, Lake Miwok, Nuuchahnulth, and Kapingamarangi also use terms for 'tree' for 'shrub, bush.'

There are also six languages, namely Ngambay, Khalkha, Upper Chehalis, Nez Perce, Yaqui, and Hawaiian, in which 'tree' and a configuration of trees, that is, 'forest,' are colexified. Furthermore, Nunggubuyu and Lengua colexify 'tree' with 'canoe,' and in Hawaiian *lā'au* also denotes the "canoe endpiece." Maxakalí and Tsafiki colexify 'tree' and 'bone,' and in Kiliwa, *t-haq=tay* is analyzable as 'OBJ-bone=be.large' and also means 'arm, limb.'

In Rendille, Ngaanyatjarra, and Yir Yoront, terms for 'tree' are also used for 'thing' in general. Similar parallelism is found in Ngaanyatjarra and Yanomámi: in both languages, relevant terms also refer to a wooden artifact in general (a pattern also occurring in Rotuman, while the relevant Tetun term, similarly, can refer to a 'tool' or 'instrument' generally when occurring in compounds), as well as to a 'gun' or 'rifle' respectively.

Other associations include: Hausa *bishiya* is also the name of a "kind of metal helmet worn by warriors" inter alia, and Khoekhoe *haii* also means 'marihuana' in informal language. Ngambay *kake* also means 'place,' Rendille *géey* also 'dance, song,' and Muna *pughu* is also used with the meaning 'source, upholder.' Ngaanyatjarra *ngarna* also means

“entrance to honey ants’ hole” and acts as a particle meaning “it’s only because, it’s only when” and as an adverb meaning ‘temporarily;’ *purnu* likewise functions as a conjunction. *Sko rí* also means ‘scales of fish.’ Yir Yoront *yoq* also denotes the ‘Rainbow Serpent’ and ‘cyclone,’ and *yulh*, a register-specific term, also means ‘tobacco.’ Badaga *mora* also denotes the “black-bark tree” specifically as well as a ‘winnow, winnowing fan,’ Basque *zuhaitz* can also refer to a ‘tree’ in Computer Science (this term is etymologically related to *zur* ‘wood.’ The other component is the name for the ‘oak;’ this etymology is similar to that proposed for an Indo-European term for ‘tree’). Japanese *ki* (due to borrowing from Chinese) also means ‘life-spirit’ and ‘intention,’ Khalkha *modu(n)* also denotes the ‘domino’ game, and Laz colexifies ‘tree’ with ‘milk’ (accidentally, due to phonemic merger in the dialect from which the data come). Dongolese Nubian *ǵóww(i)* denotes the ‘black tree’ (*Acacia Arabica*) specifically and Upper Chehalis *yámč* the ‘Douglas Fir’ specifically, *łíšałš* in the latter language also means ‘pipestem,’ and Cheyenne *hoohtsêstse* is more narrowly the name of the ‘Cottonwood tree’ (see Trager 1939 on this pattern). Nuuchahnulth *łaqafas* contains *łaq-* ‘to grow.’ Tuscarora *urę?eh* also means ‘shaft of a cart,’ while Central Yup’ik *napa* as a verb means ‘to stand upright’ and *uqvaraq ~ uqvik ~ uqvíaq ~ uqvigaq* also denotes the ‘willow’ specifically and is said to contain the “deep root” *uq-* ‘shelter.’ Cavineña colexifies ‘foam,’ and Guaraní *yvyra* contains *yvy* ‘earth.’ Jarawara *awa* is also used with the meaning ‘garden’ and in addition is the name of “a house made of boards in the sky where spirits are said to dwell.” Kaingang *ka* also means ‘mosquito,’ and Miskito *dus* is also used with the meaning ‘rheumatism.’ Rama *kát ~ káat ~ ikát* also means ‘foot,’ and Wayampi *wila* also ‘bird, dance consecrated to birds.’ Bislama *tri* also means ‘three’ (due to collapse of Engl. *tree* and *three*), *wud* also ‘carving,’ and, archaically, ‘penis,’ and *hed* also ‘head’ and ‘bow of ship.’ Fijian *kau* also means ‘to carry’ inter alia, and Bwe Karen *θo* also means ‘head-louse’ inter alia. Hani colexifies ‘tree’ with stalk,’ Hawaiian *lā’au* also may refer to a ‘club,’ ‘picture frame,’ ‘medicine’ and conveys the meaning of ‘hardness, firmness,’ and presumably from there on also may refer to a “lump or knot in flesh,” ‘cramp,’ and ‘male erection.’ Samoan *lā’au* also denotes an ‘apparatus’ and ‘machine.’ Rotuma *’qi* colexifies ‘tree, plant’ with ‘stiff, rigid’ and other meanings, *hū* inter alia means ‘lower end,’ Takia *ai* also means ‘pelvis,’ and Tetun *hun* also ‘bottom, base’ and ‘beginning, origin.’

#### 66. *The Valley*

Representation in Database: 77%

Motivated: 47.0%

Thereof Analyzable: 20.2%

Thereof Colexifying: 27.8%

Thereof by Contiguity: 19.1%

Thereof by Similarity: 6.4%

Recurrent associated meanings: gully/furrow/ditch/gorge/canal, plain/low

land, river, flat/flat land, field/meadow/lawn, river bed, cave, clearing,  
mountain, open/opening, pampa, hole, prairie, stomach, water, bay, water  
course, descent

Motivated terms for ‘valley’ are mostly contiguity-based, although there is at least one clear recurrent pattern with metaphorical transfer realized by analyzable terms, namely

that of 'stomach' to 'valley': Khoekhoe has *!goa-!nā-b* 'depression/ditch/gully-stomach/interior-3SG.MASC,' Burarra *mu-gochila* 'CLASS.DOMESTIC-abdomen' (the term also means 'depression in the ground generally'), and Miskito *il byara* 'water abdomen.'

It is relatively rare to feature a term for 'valley' with a constituent meaning 'water' cross-linguistically. Wayampi has *iapi* /ii-api/ 'water/river source' colexifying 'spring' and 'dew,' Hawaiian *kaha-wai* 'place-water' for both 'river' and 'valley,' Bwe Karen *lò do* 'water flat.land,' and Lenakel *napinu*, analyzable as /napin-nu/ 'drain/gutter-water' (a semianalyzable term is found in Yuki). There are several languages in the sample in which 'valley' is associated lexically with the body of water creating valleys in the first place, namely the 'river.' 'River' or 'creek' and 'valley' are colexified in Khalkha, Nez Perce, Bora, Huambisa, and Hawaiian. Badaga *a:ru* assumes the meaning 'river' in toponyms (and otherwise also means 'to jump,' 'to get cool,' and 'six'), and a semianalyzable term of this kind is found in Kwoma. Similarly, Dongolese Nubian, Abzakh Adyghe, and Rotuman colexify 'watercourse.' Colexification of 'river' and 'valley' is also found in Arabic (Wehr 1976: 1059). Furthermore, Central Yup'ik (Nunivak Island dialect) *kuigyaneq* perhaps contains *kuig* 'river' and the postbase (see § 4.4.2.) *-neq* 'area of,' while the Lake Iliamna dialect has *kuig-na-yuq* 'river-??-thing.like.' Hani *lolxaq* contains the classifier for rivers *lol* and *xaq* 'cut open' (there is another term involving *lol*, *lolgov*, with *gov* denoting a 'thin emaciated person or animal'), One *foli sila yarole yol* appears to contain *foli* 'water, river' and *yolu* 'fall,' and Miskito *awala bak plapi tasbaya nani* contains *awala* 'river,' *bak* 'through' and *tasba* 'land.' Sentani, Abzakh Adyghe, Khalkha, Pawnee, Lengua, and Rotuman colexify 'valley' and 'river bed' (the Khalkha term is also rarely used with the meaning 'large lake' *inter alia*). As in Hani, terms in many sampled languages betray an association to 'open' land of some sort. Bororo has *boe-ia* 'thing-opening.' Similarly, Cahuilla *pánuwenik* contains *-pánuwen-* "to flare out, to spread out in the full," and Kashaya *?ama kiya-qalli* is perhaps analyzable as /?ama kiya-qali/ 'earth extend-wide.' In Buli, Ineseño Chumash, San Mateo del Mar Huave (by the analyzable term *najmiüc iüt* /na-ajmiüc iüt/ 'AGT-fall land'), Nez Perce, Santiago Mexquititlan Otomí, Copainalá Zoque, Arabela, Embera, Kaingang, and Fijian, 'valley' is colexified with 'plain' or 'low ground, low land,' in Khalkha (by a semianalyzable term containing a verb meaning 'to blink, chop'), Chickasaw, Itzaj, Lake Miwok, and Pawnee (by a term literally translatable as "flat ground place") with 'clearing,' and in Upper Chehalis, Chickasaw, and Nez Perce with 'prairie,' in Chickasaw with 'desert,' in Bezhta with 'field' and 'village square,' in Kolyma Yukaghir with 'tundra,' and in Arabela, Cashinahua, and Tehuelche with 'pampa.' Commonly, 'valley' is also colexified with 'flat' or 'flat land,' as in Badaga, Santiago Mexquititlan Otomí, Wintu, Embera, and Kaingang, or is expressed by complex terms featuring elements with this meaning, as in the Pawnee term just mentioned, Yaqui *pa'a-la bwia* 'plain-ADJVZ land,' and Bwe Karen *lò do* 'water flat.land' (a semianalyzable term with a constituent meaning 'flat' is furthermore found in Yuki), and in Basque, Bezhta, Lake Miwok, Nez Perce, Wappo, and Wintu, 'valley' is colexified with 'field,' 'meadow' or 'lawn' (see also § 6.3. and Buck 1949: 28 for this connection in Celtic and between cognates in Latvian and Ancient Greek). Analogously, Maxakalí *hāpxa-hit* is analyzable as 'field-remain.' Five sampled languages colexify 'valley' with 'cave.' These are Ngambay (also with 'rapidly'), Khalkha, Nivkh, Nez Perce, and Pipil.

Similarly, Buli, Kwoma and Nuuchahnulth colexify ‘hole’ (also “depression in ground” and ‘drains’ in Buli and “large hole or depression in the ground” more specifically in Kwoma). Note also that Blackfoot *sstsikómm*, which also denotes a ‘coulee,’ contains the verb *sstsikkii-* ‘be hollow’ (compare evidence from Ancient Greek reported in Buck 1949: 28 for this connection). Buli, Hausa, Baruya, Burarra, Kwoma, Kosarek Yale, Badaga, Nivkh, Khalkha, Itzaj, Nez Perce, Oneida, Wappo, Bwe Karen, Fijian, Hani, Hawaiian, Lenakel, and Rotuman colexify ‘valley’ with ‘gully,’ ‘furrow,’ ‘ditch,’ ‘gorge,’ or ‘canal’ (compare the Khoekhoe term *!goa-!nā-b* ‘depression/ditch/gully-stomach/interior-3SG.MASC’ mentioned above).

Two languages, Hani and Samoan, have terms for ‘valley’ that might contain elements meaning ‘to separate’ or ‘be divided.’ In Efik, the relevant term has a constituent meaning ‘to descend, descent,’ and these meanings are colexified in Huambisa. There are also two languages in the sample which, by spatial contiguity, colexify ‘valley’ with ‘mountain,’ namely Buin and Tsafiki. Also, there are complex terms on the basis of ‘mountain’: Yoruba has *ilẹ̀ lārin òkè méjì* ‘earth between mountain two,’ Carrier has *dzel-î-krez* ‘mountains-REL-between,’ and Tetun *foho-leet* ‘mountain-space’ (compare the colexification of ‘valley’ and “interval, space between” in Piro). In addition, a semianalyzable term where one of the constituents is ‘mountain’ is found in Kwoma and Sedang. Takia *ilo-* also means ‘inside, emotions’ and ‘bay,’ and the latter meaning is also colexified in Hawaiian, where the relevant term is restricted to place names, and the Central Yup’ik (Yukon dialect) term *ilutak* contains *ilu* ‘interior, area inside; inner feelings.’

Other associations include: Buli colexifies ‘valley’ with ‘swamp,’ Hausa *kwari* also means ‘quiver,’ and, in impolite usage, “an infant being.” The Katcha term *thare (na)kidhanε* is analyzable as ‘moon middle,’ Kaluli *gagodo:* may also refer to a ‘dip between ridges,’ and Kwoma *tabotii waw*, which can also refer to a “deep pit,” contains *waw* ‘deep.’ Kyaka *anda*, denoting an “open valley area,” may also refer to a ‘hut’ or ‘house’ inter alia. Muna *labhanga* colexifies ‘vacancy, opening’ as well as “place where animals frequently pass through,” and *solobhangka* ‘shallow valley’ appears to contain *solo* ‘current, flow’ and *bhangka* ‘boat, canoe.’ Sentani *jaba* might contain *ja* ‘to sink,’ while Kosarek Yale *na’ob* also means ‘pulse, pulse area’ and, register-specifically, ‘near.’ The Abzakh Adyghe term *legʷane* contains *legʷ* ‘base’ and *ne* ‘eye.’ *Tʷaʷe* is analyzable as /tʷe-ʷe/ ‘twice-sink.in.’ Badaga *taggu* also means “to stop, stanch, staunch, thin down the flow,” and “to lower oneself, humble oneself,” while Basque *bailara* colexifies “watered meadow” and also means ‘borough, quarter.’ Chickasaw *kochchaafokka* is analyzable as /kochchaʷ aa-fokha-/ ‘outside-LOC-be.in-NMLZ,’ and *yaakniʷ hayakaʷ* is literally ‘earth way.off.somewhere’ and denotes an ‘open place’ in general. Haida colexifies ‘valley’ with ‘slough’ and “insides surface or area, insides” (note that *ilutak*, a term found in the Yukon dialect of Central Yup’ik, contains *ilu* ‘area inside’). The Kiliwa term *?+mat=xuʷsawy* is analyzable as ‘DN+earth/land=clean/clear’ (with *xuʷsawy* being itself morphologically complex). Lesser Antillean Creole French *valé* also means ‘to swallow, drink’ inter alia (due to collapse of Fr. *vallée* with *avalér*?), while Pawnee *huukiihaar* is literally “enclosure on a surface.” San Lucas Quiaviní Zapotec *baʷi* also means ‘really, actually’ inter alia (the reading as ‘valley’ is < Span. *valle*), and Guaraní *yvytyrokái* and *yvytypaʷû* contain *yvy* ‘earth.’ Lengua *mitmegyag*

might be related to *mitmeyi* “to dig, to scratch up,” and Ancash Quechua *raqra* also means ‘crack, split.’ Fijian *qākilo* seems to contain *kilo* “a low place, ravine, hollow in centre of whirlpool or water being stirred in a cup” (also meaning ‘low, hollow, depressed’ adjectivally), and *qiloqilo* is reduplicated from *qilo* ‘hollow in tree where water stagnates.’ Hawaiian *awaawa* ~ *awāwa* ‘valley, gulch, ravine’ appears to be reduplicated from *awa*, meaning ‘harbor, cove’ and ‘channel, passage through reef.’ *Kuawa*, a poetic term for ‘valley’ in the same language also means ‘guava’ due to English influence, and Malagasy *lohasàha* is analyzable as ‘head-rice.field.’

#### 67. *The Volcano*

Representation: 20.95%

Motivated: 52.67%

Thereof Analyzable: 97.97%      Thereof Colexifying: 2.03%

Thereof by Contiguity: 70.92%      Thereof by Similarity: 26.02%

Recurrent associated meanings: burn/fire, mountain/hill, earth/land, melt, explode

Frequently, complex terms for the ‘volcano’ are of the lexical type and minimally consist of elements meaning ‘mountain, hill’ and ‘fire, fiery’ or ‘burn,’ as in Buin *menu oguai* ‘mountain fire,’ and at times also further constituents, as in Nivkh *t’uyr-kir t’a bal* ‘fire-INSTRUMENTAL breathe mountain.’ Such terms are also found in Yoruba, Kyaka (which also has an alternative term where ‘lightning’ instead of ‘fire’ figures as one of the other elements alongside ‘mountain’), Basque, Khalkha, Welsh (where *llosg* means, alongside ‘burning,’ also “arson, inflammation, scald”), Xicotepec de Juárez Totonac, Guaraní, Miskito, and Samoan.<sup>7</sup> Alternatively, in some languages also ‘land’ or ‘earth’ rather than ‘mountain’ is featured in complex terms as in Efik, where *ayua obüt ikañ* contains *obüt* ‘land’ and *ikañ* ‘fire,’ Guaraní *vyv-rata* ‘earth-fire’ (Guaraní also has *tata-po* ‘fire-hand,’ *tata-sê* ‘fire-come.out, and *tata gue’êha*, containing *tata* ‘fire’ and *gue* ‘to extinguish’), Bora, where *tsátsihdyu íñúji pañétú cúújuwa íjchívyéne* contains *íñúji* ‘earth’ and *cúújuwa* ‘fire,’ Hawaiian *ahi* ‘ai honua’ ‘fire consume land,’ and Tetun *rai-suut* ‘earth-gap.’ Another complex term involving ‘fire’ is found in Baruya, which has *wa’ni-dika* ‘permanent-fire’ (this term indeed also denotes a “fire that has been burning for a long time”).

Moreover, in two sampled languages, Chickasaw and Chayahuita, terms for ‘volcano’ contain an element meaning ‘to melt’: *tobaksi’ bila-* ‘coal melt-NMLZ’ (this term also denotes ‘lava’), and *na’pi soquirin-so’ no’pa quëran pashí tēnin-so’* ‘stone melt-3SG.SUB earth from ?? say-3SG.SUB’ respectively. Also in two sampled languages, relevant terms contain a constituent meaning ‘to explode’: Wintu *p<sup>h</sup>uyuq p<sup>h</sup>uqa* ‘mountain explode/erupt’ and Arabela *jija taaniu* ‘earth explode.’

Other lexico-semantic associations are few: Itzaj has *qj-noj witz* ‘MASC-big mountain’ (this term is glossed also as ‘mountain’ itself), and Hawaiian *pele* also means ‘eruption,’ ‘lava flow,’ and ‘soft, swollen, fat,’ and is also the name of a volcano goddess.

<sup>7</sup> Japanese has this pattern, too, but the consulted source does not mention the relevant term and hence this case is outside the present sample.

68. *The Waterfall*

Representation: 67%

Motivated: 52.0%

Thereof Analyzable: 38.4%

Thereof Colexifying: 13.6%

Thereof by Contiguity: 31.5%

Thereof by Similarity: 6.6%

Recurrent associated meanings: water, current/rapids/cataract, fall, cliff/precipice, river, jump, flow, down

English *waterfall* is a representative of the cross-linguistically most common structure of analyzable terms for the ‘waterfall,’ featuring constituents meaning ‘water’ and ‘fall.’ In the sample, they are found in Khoekhoe, Noni, Basque, Welsh (in addition, there is another term which directly colexifies ‘waterfall’ with ‘flow, fall’ in this language, and Wintu *tEk* likewise colexifies ‘waterfall’ with ‘flow,’ next to ‘extrude, be extruded’ and ‘press’), Chickasaw, San Mateo del Mar Huave, Itzaj (colexifying ‘jet of water’), Kashaya, Santiago Mexquititlan Otomí, Yaqui, and Sáliba. Moreover, for Haida *quuga*, compare *quugaa* ‘have fall.’ The source notes: “*quuga* may in fact have originated as a nominalization of the stative [verb *quugaa*] rather than the latter being derived from a noun.” There are also many languages with terms with a similar structure, but where the semantics of the constituents is slightly divergent. Yoruba has *òṣòrò omi* ‘pour.down water’ (note in this context that Buin *parukuna* is derived from a verb meaning ‘to spout, to pour’), Sora *tũsar'da:n*, containing *t'ũ:ŋ*- ‘to collapse, drop down, pierce’ and *d'a:-* ‘water,’ Oneida *tetwa?śátha?*, analyzable as /te-w-a?saht-ha?/ ‘DUALIC-NEUT.AGENT-drop:CISLOCATIVE-HAB,’ Pawnee *kictakaahak*, literally “water to pass down,” which is a verb meaning ‘be a waterfall’ as well as ‘water to drip’ and ‘be a rapid’ and contains elements meaning ‘be liquid’ and ‘down,’ Xicotepec de Juárez Totonac *yujyā xcān*, presumably /yujā xcān/ ‘move.down water,’ Yana *ba-ri?mau-xa* ‘spill/flow-place-water’ (this term also denotes a particular site; note also that in Carrier there is a semianalyzable term for ‘waterfall’ containing a verb ‘to flow’ and compare also Cheyenne *anôhehééóó'e*, which is analyzable as /anôhe-hée'o'tsé/ ‘down-spill’ and also means ‘to pour down’ when interpreted verbally), Takia has *you i-skalik da* ‘water 3SG-pour.away IPFV,’ Tetun *beetudak* containing *bee* ‘water’ and *tuda* ‘hurl, fling,’ and Tuscarora has *yuhtawé?e*, containing *-htaw-* ‘stream of water, current of water’ and *-e-* ‘fall.’ Chukchi has *emlaratγəγən*, containing *eret-* ‘fall’ (subject to vowel harmony), and Rama *kála taik* ‘fallen nose.’

As in Tuscarora, there are also terms in which ‘river’ rather than ‘water’ alongside ‘fall’ figures. These include Kapingamarangi *monowai doo* ‘river fall’ and White Hmong *dej poob tsag* ‘river fall cliff.’ Similarly, Yay has *ram<sup>6</sup> tok<sup>3</sup> taat<sup>2</sup>* ‘water fall cliff’ as well as *taat<sup>2</sup> ram<sup>6</sup>* ‘water cliff,’ where *taat<sup>2</sup>* is “cliff over which water falls,” “cliff which is the site of a waterfall” (a semianalyzable term including ‘river’ is found in Embera). In line with the Southeast Asian association with ‘cliff’ by overt terms, Bwe Karen colexifies ‘waterfall’ and ‘steep cliff, precipice,’ inter alia. Similar to the association with ‘cliff’ in the Southeast Asian languages is also Copainalá Zoque *na'atɔŋgomɔ'nguy* /na'-tɔŋgomɔ'nguy/ ‘water-precipice.’

In Yoruba, Hani, and Hawaiian, there are terms for ‘waterfall’ containing verbs referring to some sort of jumping motion: Yoruba has *ì-takiti omi* ‘NMLZ-to.somersault water,’ Hani (*wulquvq*) *coq-ceiv* ‘(water) jump-break’ alongside *wulquvq coq-ciivq* ‘water jump-pinch/choke,’ and Hawaiian *wai-lele* ‘water-jump/leap.’ Other complex terms for ‘waterfall’ with ‘water’ being one of the constituents include Meyah *mei ofoská* containing *mei* ‘water, river, sperm’ and presumably *ofos* ‘skin, mountain top,’ Sora *taŋ'surda:n /taŋ'sur-d'a:n/* ‘rise-water-N.SFX,’ Guaraní *y-tororō* ‘water-gush’ and *y-tu*, probably analyzable as ‘water-father,’ Piro *gijrotga /gijrota-giga/* ‘forehead-water,’ Wichí *inot n'oyij* ‘water way,’ and Great Andamanese *inal'ârçhâr*, containing *îna* ‘fresh water’ and *châr* ‘spring of water.’ Semianalyzable terms for ‘waterfall’ on the basis of ‘water’ are found in Berik, Kyaka, One, Yir Yoront, Nivkh, Bislama, and Lenakel. In Kiowa, ‘waterfall’ is *zout-syŋn-goup* ‘current-small-hit’ and in Ket it is *qakun /qa'-qūn/* ‘big-current.’ ‘Current,’ ‘rapids,’ or ‘cataract’ are colexified in Khoekhoe, Kwoma, Khalkha, Nivkh, Upper Chehalis, Itzaj, Pawnee, Qui-leute, Aguaruna, Bora, Chayahuita, Cubeo, Hupda, Miskito, Ancash Quechua, Yanomámi, Tetun, and Hawaiian.

Other associations include: Ngambay *kóró* also means ‘storm’ and denotes a kind of snake, and Swahili *maporomoko* consists of the verb *poromoko* “to slide down in a mass” and the noun class prefix *ma-*. Kyaka *jingi* also means ‘flower, bloom.’ Muna *kasapa* is derived from the verb *sapa* ‘to splash,’ and Rotokas colexifies ‘waterfall’ and ‘steep slope’ as well as ‘stone face.’ Kosarek Yale *modun* ‘place of a waterfall’ also means “upper edge of a landslide.” The Bezhta term *lišiyo* contains a verb meaning ‘to entangle’ and the past perfective participle suffix. Blackfoot *ohpsskonaka'si* also means “to spurt/flow rapidly” as well as ‘geyser.’ Nez Perce *tíkem* also means ‘dam’ and ‘fish bladder.’ Nuuchahnulth *tuxʷił* also can refer to the ‘Sproat Lake falls’ and “the falls on the Sarita river,” and Oneida *tetwaʔsáthaʔ* also to the ‘Niagara falls’ specifically. Central Yup'ik *qurrlugtaq* contains *qurre-* ‘to urinate; to spawn (of fish).’ Bora *waapéwa* colexifies ‘spout, stream’ and contains *waapé-* ‘to drip’ and the classifier *-wa* for table-like entities, and *nééwabya* contains *nééwa* ‘stone’ and the classifier for three-dimensional objects *-bya*, while Miskito colexifies ‘waterfall’ with ‘rock, big pebble in the middle of river.’ Bwe Karen *bla* also means ‘to wash the face,’ and Samoan *āfu* also means ‘(to) sweat,’ ‘(be) heated,’ and ‘to wither.’

#### 69. The Wave

Representation: 70%

Motivated: 19.8%

Thereof Analyzable: 17.5%

Thereof Colexifying: 2.3%

Thereof by Contiguity: 2.9%

Thereof by Similarity: 14.9%

Recurrent associated meanings: water, river, swell/swelling, sea, wave (techn.),  
skin/surface, stir, gust, surge

Motivated terms for the ‘wave’ (‘billow,’ ‘ripple,’ ‘breaker,’ ‘surf’) are relatively rare cross-linguistically. Where they occur, they are likely to be similarity-based. In many languages, terms feature verbs referring to the motion of the waves, a situation also found in Indo-European languages (Buck 1949: 40). Wayampi *palanayaapiyapi* is analyzable as /palana-y-

apiyapi/ 'sea-NMLZ/REFLX-to.surge,' and Abzakh Adyghe *we- $\lambda$ e* as 'surge.up-foot.' Baruya *mavajinaaka* colexifies 'wave' with 'swirl,' Muna *kaendo* is derived from *endo* 'to shake' and *lolebhata* "calm waves without foam" might be related to *lole* 'to roll around.' Kosarek Yale *mak kola'ebna* contains *mak* 'water' and a verb meaning "to make stirring movements" (not necessarily of water, though), and Hawaiian *nalu* ~ *nanu* colexifies 'wave' and 'to stir' inter alia. Kyaka *ipwua* ~ *ipyua kaso kole minyingi* contains *ipwua* ~ *ipyua* 'water' and *kasu kole* 'up and down,' (with *kasu kole minyingi* glossed as 'wave movement'), Biloxi has *ani' xoxo-ni' /ani' xoxo'-ni' /* 'water swing-CAUS,' Cheyenne *-nêšêške'sevó* contains *nêšê* 'to wash' and '*sevó* 'to flow,' a constituent of San Mateo del Mar Huave *ateapteapeay* is *ateapte-* 'to move by wind,' Santiago Mexquititlan Otomí *munts'i* is analyzable as /m-punts'i/ 'NMLZ-to.relapse,' Yaqui *bawero'aktim* as /bawe-roa-k-te-im/ 'sea-turn-PFV-INTR-PL,' Wintu *le(.)w* is connected to a root meaning 'vibrate, oscillate,' Bororo *pobo amagadu* is analyzable as 'water agitation,' and Guaraní *y-pu'ã* is analyzable as 'water-rise.up.' Abipón *lilikatka* contains *ili(k)-* 'movement of water,' and finally, Dongolese Nubian *árre* goes back to \**ánre*, containing *án-* 'to go.'

However, there is also a wealth of terms for 'wave' in the sample that make reference to the shape rather than the movement of waves, and these are frequently metaphorical in nature. Specifically, there are a lot of body-part metaphors (as already seen in the Abzakh Adyghe term mentioned above): Yir Yoront has *yuwl-man* 'sea-neck,' Carrier *tha-tši* 'water-head,' and Bororo *iagajaga* and *iagiri* seem to contain *ia* 'mouth.' Other metaphorical denominations include: Hausa *rak'umi-n ruwa* contains *rak'umi* 'camel' and *ruwa* 'water' (*rak'umi* can also refer to the "blossom of locust-bean tree before it is fully out"), Kanuri *kàzáà njî-bè* is analyzable as 'spear water-of,' Bezhta colexifies 'wave' with 'horse,' and Pawnee *kicpiirakus* is analyzable as /kic-wiira-kus/ 'be.liquid-upright-be.sitting.' The Piro verb *hawokhata* 'to be turbulent, have waves' contains *hawoka* 'to blow,' and *špurha* contains *špu* 'edge.' Yay *ram<sup>6</sup> for<sup>4</sup>* is presumably analyzable as 'water roof,' and *me<sup>5</sup> for<sup>4</sup>* as 'mother roof.' Guaraní also has *y apeno* containing *y* 'water' and *ape* 'skin, surface,' and the latter two meanings are associated with 'wave' by direct colexification also in Welsh, while there is a semianalyzable term in Central Yup'ik.

As seen from the data discussed so far, frequently complex terms for 'wave' unsurprisingly contain a constituent meaning 'water.' Further terms of this structure are Kwoma *uku veereveer* 'water breeze/light.wind,' Bororo *pobu-to* 'water-inside,' and Guaraní *y joapy* 'water connected.' Rama *albríni* in *sí albrínima* 'wave' means "waving, grinding" (*sí* is 'water'). Semianalyzable terms with 'water' are found in Kwoma, Kosarek Yale, Cavineña, and Guaraní, while Ngambay directly colexifies 'wave' with 'water' as well as 'river' and 'well,' Sko and Lenakel have semianalyzable terms where one of the constituents can be identified to mean 'sea.'

Furthermore, Hawaiian '*ale* rarely means 'gust' inter alia, and White Hmong has the analogous complex term *nthwv dej* 'gust water.' In Efik, the 'wave' is *mbufut akpa* 'swelling river,' for which compare Berik *fo buk* 'wave' with *bukna* 'to swell' (*fo* is 'water, river, lake'); Basque colexifies the meanings, and the association is also present in Tasmanian according to Plomley (1976: 380) and in Indo-European languages (evidenced for instance



in Latin and Ancient Greek, Buck 1949: 40). In Basque and Mandarin, the relevant terms can inter alia also refer to a 'wave' in the technical sense, e.g. a 'radio wave.'

Other associations include: Rendille colexifies 'wave' with 'storm' on a lake or the sea, Swahili with "bulrush millet, eleusine," and Kwoma *uku woya* also denotes a "decorative zigzag band on clay pots and other artefacts." Kyaka *pyakao pulyuo lupyuo pingi* contains *pulyuo* 'upwards' and *lupyuo* 'down, descending.' Toaripi *aroaro* also denotes 'charcoal,' probably by extension from there "the Frigate Bird; black with white breast," as well as a 'squid' and the *Mimosa pudica*. Sahu *moku* also means 'dizzy.' Basque *itsaski* contains *itsas* 'sea' and also denotes 'seafood,' while *olatu* as a verb means 'to tame, domesticate' and 'to beat, pound.' Khalkha *kyrkire*, as a verb, also means 'to growl, grunt, snarl,' *calgija* contains the verb *calgi-* 'to splash' and *dolgija(n)* is related to *dolgi-*, meaning 'to wave, undulate,' but also "to be restless, tempestuous, irascible." Welsh *ton* also means 'lay-land' and 'broken.' Blackfoot *ohpai'kímsskaa* may contain *ohpai'piyi* 'to jump,' Kashaya *daluw* also means "rub or spread something by hand," Chickasaw *bo'kalhchi* 'to come in waves' also means "to be beaten (of eggs), to be splashed," Lesser Antillean Creole French *lanm* also means 'blade,' Itzaj *kukul* is a verbal noun derived from *kul* 'to roll,' and Nuuchahnulth *kʷax* also means 'spray.' Central Yup'ik *yuulraaq* also denotes a 'thin flexible sheet of ice on ocean.' Bora *tyocáhco* also denotes the sound of water splashing against something, Cayapa *sela* is also the term for the 'sky.' Hani *eeldaol*, *eeldaol bi*, and *eeldaol hhev* contain *daol*, meaning 'to exist, lead the way' and 'one line' inter alia; *eel* is glossed as 'to laugh, to smile,' but occurs in many complex terms for water-related meanings. Quaintly, *hhev* in the last variant means 'to wave one's hand, to gesture.' Fijian *ua* 'wave, tide' also denotes 'veins, tendons' and 'muscles.' Hawaiian *nalū ~ nanu* also inter alia means 'to ponder, to reflect' and '*ale* also 'for tears to well in the eye.' Malagasy *àlona* also means 'jealousy,' Rotuman *vālu* also 'eight,' Samoan *galu*, as a verb, also 'for the sea to be rough,' and White Hmong *twv* also 'to dare' inter alia.

#### 70. The Wax

Representation: 55%

Motivated: 45.7%

Thereof Analyzable: 24.5%

Thereof Colexifying: 21.2%

Thereof by Contiguity: 15.2%

Thereof by Similarity: 25.5%

Recurrent associated meanings: candle, honey, bee, resin, fat/grease, faeces, tar, ear-wax, house

Many denominations for the 'wax' are metaphorical in nature cross-linguistically, with or without 'bee' acting as contiguity anchor. Most frequent, however, is a pattern of colexification by provenience contiguity, namely that with 'candle,' found in Khoekhoe, Dongoles Nubian, Muna, Greek, Khalkha (the relevant term is a borrowing from Chinese in this sense; it has also other indigenous meanings), Welsh, Highland Chontal, Itzaj, Xicotepec de Juárez Totonac, and Tetun (in the latter languages, from 'candle' extended also to 'lamp'). Furthermore, Chayahuitta has *pa'nan*, where *pana* is 'side of the candle' and *-nan* an instrumental suffix. Taken together, however, metaphor-based associations with substances

which are, like ‘wax,’ semi-solid, formable and/or have a similar texture to the target concept are more frequent. For instance, the Khoekhoe term exhibiting this pattern of colexification, *tnuru-llnui-b*, is analyzable as ‘knead-fat/oil-3SG.MASC.’ In fact, ‘fat’ and ‘wax’ are colexified in Samoan (alongside “medicinal cream, wax”), and Imbabura Quechua *mishki wira* is analyzable as ‘bee grease,’ while Abzakh Adyghe *sef* may contain *se* ‘fat’ and *fə* ‘white.’ Further, Efik has *a’dan u’tōñ* ‘oil grease ear’ (compare *utōn unam* ‘glue’ “[s]o called from being supposed to be made of the ears of animals”), and there is a semianalyzable term where a constituent may mean ‘greasy’ in Hawaiian. In Tuscarora, Hawaiian, and Tetun, ‘wax’ and ‘resin’ are colexified (Tuscarora also colexifies ‘cement,’ ‘glue,’ ‘gum,’ ‘jelly,’ and ‘syrup’ by the same term, and Tetun also ‘candlenut’ and ‘to prune, clip’), while in Wintu the association is realized by the analyzable term *hub-in ceki* ‘bee-LOC pitch.resin.’ Similarly, Haida colexifies ‘wax’ with ‘pitch, tar’ (and also ‘chewing gum’), Tuscarora also colexifies ‘wax,’ ‘tar,’ and other meanings, and Kildin Saami has *veazvušk-tārr’v* ‘wasp-tar.’ In Buin, ‘wax’ and ‘faeces’ (and also ‘fine powder’ generally) are colexified, and this association is mirrored by complex terms in Buli (*si-beung* ‘fill.up-faeces/droppings’), Carib (*wano weti* ‘bee excrement’), and Tsafiki (*chiná pe* ‘bee excrement; compare Cayapa *chiñape*, where the general term for ‘bee’ is, however, *tanda*, and *chiñilla* is a term for a type of bee similar in appearance to a wasp). There are also many languages in which ‘wax’ is associated with ‘honey.’ Herein, Buli is the only language with colexification (though Nungubuyu has a broad term that may include reference to ‘wax’ as well, although not explicitly glossed so), while in other languages the association is realized formally by complex terms: Kanuri has *kàmàgàn-mí* ‘honey-son.of,’ Kiliwa *mi?yawy=smaq* ‘honey=leaf,’ Guaraní *eiraity* /eíra-ty/ ‘honey-urine/juice,’ Hupda *něŋ cak* ‘honey mass,’ Toba *pic l-apa* ‘honey 3SG.POSS-wasp.nest/honeycomb,’ and Great Andamanese *ája-pij* ‘honey-hair/feather.’ Miskito and Rama colexify ‘wax’ with ‘ear wax’ (Miskito also with ‘scab’ and ‘scabies’). Semianalyzable terms involving ‘honey’ are featured in Hausa (where ‘honey’ is colexified with ‘bee’), Guaraní, and Hani. Moreover, two languages, Yaqui and Fijian, have terms with a metaphorical transfer from ‘house’ to ‘wax’: *muumu jo’ara* ‘bee house’ and *drega ni vale ni vī*, containing *drega* ‘gum’ and *vale* ‘house’ respectively. Other complex terms with ‘bee’ are Chickasaw *fohi’ bila-’* ‘bee/honey melt-NMLZ’ and Tetun *bani-isin* ‘bee-flesh,’ and semianalyzable terms with ‘bee’ are found in Hausa, Katcha and Huambisa (for the unknown element *nujiri* in this language, compare *nujin* ‘egg?’).

Hausa *kaki* also means ‘mucus’ inter alia, Muna *lili* also to “go or travel around,” and Basque *zira* also ‘shoe polish.’ Noni *kelay* is formed by adding the noun class prefix *ke-* to *lay*, which means ‘to seal, glue.’ Gurindji *tarla* also denotes “wax” from the Spinifex plant, Badaga *mekku* also means ‘cud,’ Lesser Antillean Creole French *si* also “sour, having an acid taste” inter alia, and Pipil (Cuisnahuat dialect) *chapah* is also the name of a game played with wax. Aguaruna *dují ~ nují* also means ‘nose, beak’ and ‘prow of canoe,’ Bororo *bori* also denotes the bony scales of an alligator, and Sedang *pet* also means ‘to plant’ inter alia.

71. *The Whirlpool*

Representation: 55%

Thereof Motivated: 50.1%

Thereof Analyzable: 43.1%

Thereof Colexifying: 7.4%

Thereof by Contiguity: 29.2%

Thereof by Similarity: 19.2%

Recurrent associated meanings: water, go around/spin/turn/twist, whirlwind, eye, vessel, current, sea

Most frequently, terms for 'whirlpool' ('eddy,' 'vortex,' 'maelstrom') are morphologically complex in the languages of the sample, containing terms for 'water' or specific bodies of water and verbs for motions such as 'to go around,' 'to spin,' 'to turn,' and 'to twist,' as in Koyraboro Senni *hari-windi* 'water-go.around/go.in.circles.' Other languages with terms with such a structure and varying water-related contiguity anchors are Bezhta, Ket (which also has the alternative term *telyil* /tel-kil/ 'mammoth-spin,' the analysis of which is unsure however according to lexicographers), Welsh, Ineseño Chumash, Haida, Kashaya, Kiliwa (*xa?-t s+wa-y-p=s+wa-y-p* 'water-SUBJ INST.LONG+spiral.motion-ATT-PASS=INST.LONG+spiral.motion-ATT-PASS'), Oneida, Yaqui, Guaraní, Toba (where the element conveying the motion is 'to return' more specifically), Hawaiian, Malagasy (where the same remark as for Toba applies), White Hmong, Samoan, and Vietnamese (where an additional element meaning 'current' is present, compare Biloxi *ani'xyu'hi kidu'nahí* 'current turn.around' and Haida *juu rwiihlrahl*, containing *juu* 'current,' *rwiihl* 'to move spirally' and *rahl* 'to stir'; this term also denotes a 'crow's nest at the back of head'). Similar terms of the derived type are encountered in Piro, Wayampi, and Hawaiian (by the term *mimilo*, which is formally a reduplication of *milo* 'twist, whirl' and also means 'curly, kinky' as well as "to roll, as to induce an abortion"). The relevant meanings are colexified directly in Kyaka (although the situation in this language is not entirely straightforward on the basis of the consulted source), Sora, and diachronically, such an association is also identifiable for a Kolyma Yukaghir term for 'whirlpool.'

There are, however, also metaphor-based conceptualizations of 'whirlpool.' In four sampled languages, there are complex terms in which 'eye' is transferred to 'whirlpool,' with 'water' or bodies of water acting as a contiguity anchor: Kanuri has *shim njî-bè* 'eye water-of,' One has *fola namna toma* 'water eye stone,' San Mateo del Mar Huave *oniiüg ndec* 'eye sea' for a whirlpool in the sea specifically, and Takia *you mala-n* 'water eye-3sg.' In three of the languages in the sample, the source concept is a vessel: Nez Perce *capa-hikayikáyica* 'funnel-shaped whirlpool' contains *cepé* 'to shape with hand' and a reduplicated version of *hi:kay* 'pot, cup, to make a cuplike shape,' and in Bororo, the 'whirlpool' is *aria-reu* 'pot-like.' In addition, Hupda *wăwŏ'y* colexifies 'whirlpool' with a "basket type having hourglass shape" and "gourd support (hourglass shape)." Basque (by the term *zurrunbilo*, perhaps related to *zurrun* 'pole' and *bilo* 'down, fine hair'? This term also colexifies 'bedlam,' 'confusion,' 'crowd,' and "heavy downpour"), Cubeo, and Fijian colexify 'whirlpool' with 'whirlwind.' Since the Fijian term is morphologically complex and contains *waitā* "water deep enough for a canoe at low tide," it appears that 'whirlpool' is the diachronically original referent of the term. The contrary situation holds in Santiago

Mexquititlan Otomí, where the ‘whirlpool’ is called *xedi dehe* ‘whirlwind water,’ and the same is true of Abzakh Adyghe and Imbabura Quechua.

Other complex terms with ‘water’ include Toaripi *ma elorielori oti* containing *ma* ‘water’ and *oti* ‘place,’ Kildin Saami *čāž’-čullm* ‘water-knot,’ Wintu *me’m kumu’ri*, containing *me’m* ‘water, river’ and *kum*, referring to the ‘sound of water rushing or roaring’ and also meaning ‘water turbulence,’ Bororo *buto-bo*, seemingly ‘rain-water,’ and Cavineña *ena cahuaitya* ‘water enrage.’ Semianalyzable terms where one of the constituents can be identified to mean ‘water’ are found in Highland Chontal, Aymara, and Hani.

Still other associations include the following: Efik *ēsik* is derived from the verb *sik* ‘to squirt;’ the term also denotes a “round thing drawing in towards the centre” and a “round net or fishing pot,” a pattern of colexification probably due to the fact that *sik* also means ‘to tie, draw together’ inter alia. Khoekhoe colexifies ‘whirlpool’ with ‘pot-hole’ in river and ‘gravel pit’ inter alia, while Muna *tehi tingkulu* is analyzable as ‘sea slope.down’ and also denotes the ocean itself when no land is in sight. Greek *roufēchtra* contains *rouf-* ‘to suck in, absorb,’ Khalkha *oil* also can refer to a ‘tuft of hair,’ and Kolyma Yukaghir *unu-jurugu* ‘eddy in a river’ is analyzable as ‘river-hole.’ Tuscarora *nekahtawakwá?nahč* contains *-htgw-* ‘stream of water, current of water’ and *-kwa?n-* ‘arc, curve.’ Cayapa colexifies ‘whirlpool’ with ‘grindstone,’ Cubeo *jōmeicobe* is derived from *jōmeñu* ‘to agitate’ by means of the classifying suffix *-cobe* for hole-like objects, and Guaraní *jepyvu* contains *yvu* ‘spring.’ Piro *ximlere* apparently contains *ximle* ‘to boil,’ for which compare Central Yup’ik *qalla-neq*, colexifying ‘spring’ with ‘eddy’ and analyzable as ‘be.boiling-thing.that.results.from.’ An-cash Quechua *muYu* also denotes a ‘rodeo’ or ‘walking around’ (Spanish gloss: ‘rodeo’). Samoan *auma* also denotes a ‘breaker,’ and Tetun *da-dula-k*, analyzable as ‘DERIV-winnow/sift-DERIV,’ also means ‘to sieve for grain.’

## 72. The Airplane

Representation: 65%

Motivated: 55.5%

Thereof Analyzable: 50.2%

Thereof Colexifying: 5.2%

Thereof by Contiguity: 4.7%

Thereof by Similarity: 28.8%

Recurrent associated meanings: fly, boat/canoe, bird, sky, machine, car/vehicle, air, high, metal, air, helicopter, wing, house, jump, thing, travel

Most frequently, complex terms for ‘airplane’ are derived from verbs meaning ‘to fly’ or ‘to fly around,’ with a variety of structural subtypes. Derivation proper is attested as the word-formation device in Gurindji, Ket (where the relevant term also can refer to a helicopter; the same is true of a simplex term in Sko), Carrier, Kashaya, Nuuchahnulth, Oneida, Pawnee, dialectal Central Yup’ik, Aguaruna, Guaraní, and Wichí, which has *wiy’o-taj* ‘fly-AUG.’ In two languages, additional elements are present, namely Nez Perce (*we’ke?yke?í* / *we’-ke?éy-e?í* ‘fly/run-move-AGT’) and Yanomámi (*isihami-yě-rewě* ‘up.high-fly-NMLZ’), and in another two sampled languages, there are complex terms that are somewhat akin to derivation with an element meaning ‘thing,’ namely Cheyenne (*ame’hahtôtse*, analyzable as /*ame’há-hestôtse*/, ‘fly.along-thing’) and Hupda (*wayd’ó?-teg*

'fly-thing'). In two sampled languages, there is a metaphorical transfer from 'house' to 'airplane' with 'fly' acting as a contiguity anchor. These are Rotokas (*papa-pa kepa* 'fly.through.air-DERIV house/cabin/building') and Kiliwa (?-wa?kw-i?+hiw 'DN-house=WH-DIST+fly'). In Kashaya and Yanomámi, there are additional constituents in the relevant terms meaning 'high' (*qalicat<sup>h</sup>mu?* /qali-hca-t<sup>h</sup>-mu-ʔ/ 'high-fly-PLURAL.ACT-around-ABS' and *isihami-yě-rewě* 'up.high-fly-NMLZ'); 'high' is also featured in Bora *cáámé-e-mi* 'be/become.high-belong.to-SCM.transport' and Piro *tenyapatšro*, containing *teno* 'high, tall, deep' and *ya* 'to go.' In Santiago Mexquititlan Otomí, 'metal' is present as the meaning of the second constituent of the respective 'airplane'-terms alongside 'to fly': *nsani bojä* 'fly iron/car,' while Bororo has *meriri kodureu*, containing *meriri* 'metal' alongside a verb meaning 'to fly' and White Hmong has *dav hlao* 'hawk iron.' Semianalyzable terms with 'fly' are featured in Basque, Khalkha (in varieties spoken in inner Mongolia), Lakhota, and one on the basis of a verb meaning 'to fly alone' in Toba. It is also not uncommon to transfer 'car,' or more commonly 'vehicle' in general, to 'airplane,' similar to Santiago Mexquititlan Otomí: Kanuri has *mààrá sámí-bè* 'vehicle sky-of,' Khoekhoe *toa-kuni-s* 'air-wagon-3SG.FEM,' Yoruba *òkò òfurufú* 'vehicle sky,' Sora *rua:ŋ-sagada-* 'sky-cart-,' Cashinahua *nai bapu* 'sky motor/motored vehicle,' Malagasy *fiaramanidina* /fiàra-manidina/ 'vehicle fly,' and Mandarin *fei1-ji1* 'fly-machine,' for which compare Japanese *hi-kō-ki* 'fly-go-machine' and Vietnamese *phi cơ* 'fly machine' and *máy bay* 'machine fly.' In the Americas, there are three languages, Kashaya, Aguaruna, and Huambisa, which colexify 'machine' or 'machine with motor' in general with 'airplane.' In all instances, the relevant terms are borrowed from Spanish, from *vapor* in the case of Cashinahua (and note that Chayahuita *huaporo*, also a borrowing from Spanish, colexifies 'steam' with 'airplane') and from *machina* in the rest of the languages. Kashaya has the optional complex term *ma'kina ca'dmuli*, which is tentatively analyzable as /ma'kina hca-v<sup>o</sup>d-mul=li/ 'machinery fly-along-around=INSTR.' More frequent than associations with 'car,' 'vehicle,' or 'machine' is the transfer from 'boat' to 'airplane,' realized almost exclusively (with the exception of Khalkha which has a colexifying term that is indeed parallel to English *vessel* in being also capable of referring to a receptacle or trough) by morphologically complex expressions, with either 'sky' as contiguity anchor, as in Hausa *jirgi-n sama* 'boat/train-GEN sky,' or more frequently with 'fly,' as in Rotuman *ahqi fere* 'ship fly' and also in Muna, Sko, Meyah, Nivkh, Chickasaw (where 'boat' is colexified with 'trough'), Cayapa, Fijian, and Hawaiian ('canoe' rather than 'boat' specifically in the case of Cayapa and Fijian), Rotuman, Samoan, and Yay, with 'jump' in Abzakh Adyghe *q<sup>o</sup>əħlate* /q<sup>o</sup>əħe-lete/ 'boat-jump,' and with 'air' in Bwe Karen (*kəbó-gəli* 'ship-air') and Hawaiian (*moku-ea* 'ship-air').

All of the second elements in the above terms recur in other configurations in other languages. Like Hausa, the meaning 'sky' figures in Ngaanyatjarra, where *yilkaringkatja* contains *yilkari* 'sky' and *katja* 'son' (this term also means 'airline, air services'), and Lengua *thlinga netin* 'movement sky.' In Miskito, 'airplane' is *pasara pali tauki ba* 'air jump travel DEM,' terms of the derived type involving the meaning 'fly' were already discussed above. In Guaraní, 'wing' and 'airplane' are colexified and there are further semianalyzable terms containing *pepo* 'wing.' Note also in this context Yir Yoront *minh-puth lon* 'animal-wing with,' but this is due to another metaphorical transfer pattern: *minh-*

*puth lon* in fact is the Yir Yoront term for ‘bird’ (compare section 6), and with this pattern of colexification, Yir Yoront is accompanied by Swahili, Toba, and Mali, where ‘large bird’ specifically is colexified with ‘airplane.’ Khoekhoe has *anis kunis* ‘bird wagon,’ Ngambay *yèl bò* ‘bird big,’ and One *tolla moa* is presumably analyzable as /tolla mo’a/ ‘bird mother’ (with ‘mother’ presumably conveying augmentative meaning, Matisoff 1992). Bakueri *méli mé ngoú* ~ *méli mé ngoú* contains *méli*, which denotes a species of bird, and *ngoú* ‘fever, cold.’ As for ‘bird,’ there is also a semianalyzable term in Guaraní.

There are also terms for the ‘airplane’ in which ‘air’ in fact is a constituent: alongside the terms already discussed, this is also the case in Welsh, where *awyren* contains *awyr* ‘air, sky’ alongside a singulative suffix. Semianalyzable terms including elements with this meaning also occur in Greek (although *aeroplánon* is in fact borrowed from French, which is in turn a learned neologism based on the Ancient Greek root for ‘air’), Lengua, and Manange (where ‘air’ and ‘wind’ are colexified).

Other more unusual patterns include: Cubeo *vúicū* contains the root *vui* also occurring in terms for ‘propeller,’ ‘lungs,’ and ‘dry tuber’ and a classifier for things that are round on the one side and plane on the other. Guaraní *kurusu veve* contains *kurusu* ‘cross,’ and Bislama *plen* (< Engl. *plane*) colexifies ‘wood plane.’

### 73. *The Ball*

Representation: 74%

Motivated: 31.7%

Thereof Analyzable: 12.5%

Thereof Colexifying: 19.3%

Thereof by Contiguity: 15.5%

Thereof by Similarity: 7.6%

Recurrent associated meanings: round, bullet, football, baseball, citrus, lump,  
play ball game, testicles, nut of dum-palm, bowl, kick, rubber

The most common of the few lexico-semantic associations with ‘ball’ (ignoring additional glosses such as ‘sphere, spheroid’) is that with ‘round, round thing’ (see Buck 1949: 907 for evidence from various Indo-European languages). It occurs either by morphologically complex terms, as in Meyah *móf ofóg* ‘wind round,’ One *malwa tala* ~ *maula rala* /malwa tala/ ‘citrus round.thing,’ Nez Perce *kapapkápap*, reduplicated from *kápap* ‘to be round,’ and Lengua *aksak yakyé* (*aksak* is ‘thing’ and *yakyé* likely related to the verb *yakyeyi* ‘to be round’) and directly by colexification in Buin, Rotokas (where the relevant term *kororoisia* appears to contain *koro* ‘fruit’), Khalkha (one of the relevant terms, *bømbyrceg*, seems to be related to *bømbyr* ‘drum’), Santiago Mexquititlan Otomí (where the relevant term also means ‘wheel’), Quileute, Wintu (where the relevant term also means ‘naked’), Embera, Hupda (where the relevant term also means ‘fruit’ and ‘nut’), and Piro (where, similarly, the relevant term also denotes a ‘large round fruit’). In addition, there are a number of languages in which the relevant term features a bound element conveying the meaning ‘round thing;’ these are Nuuchahnulth *hupk-imt* ‘roundish.thing-CHUNK.SHAPED.OBJECT (which also means ‘testicles,’ as does Bislama *bol*, which goes back to Engl. *ball*, alongside ‘sac’ and “soft belly of coconut crab”) and Cubeo *yajui-du* ‘game/sport-CLASS.ROUNDISH.OBJECT’ (for this term, compare Cahuilla *qáwpiʔl̥*, which is derived from

-*qáwpi*- 'to play a ball game,' such terms are also featured in Haida and Tuscarora, where it in addition colexifies 'setting sun') and Cayapa *sapuca* contains *puca* 'small round thing.'

However, various specific round things are colexified in many languages. In three languages of the Old World, Efik, Abzakh Adyghe, and Greek, 'ball' is colexified with 'bullet' (see Buck 1949: 907 on this extension in German; Adyghe also has terms colexifying 'ball' with 'cannon'). In both Hausa and Katcha, the relevant terms also denote the 'nut of the dum-palm' which is apparently used as a ball (in Hausa, also the Polo ball specifically and hence the Polo game as well as horse-racing inter alia). Bakueri *ɛwumá* also means 'orange,' and Buli *gbeli* is also the name of "the wooden cylinder of the *buuri*- or calabash fruit that are used by sheperds as a 'ball,' when they play their game of 'hockey'." Ngambay *bbiri* also means 'parcel, container,' and *ékiyáy* also means 'placenta.' The Yoruba term *ìṣu* 'ball, lump of anything' (the association with 'lump' is shared with Samoan and Tetun, which has *ai-kabuar* 'tool-circle/lump' and many Indo-European languages according to Buck 1949: 907-908) consists of the verb *ṣu* 'to mould in a round form' and the prefixal nominalizer *ì*-. Kyaka *konda* is also the term for a 'kind of yam;' the term denotes an oval-shaped ball similar in shape to the fruit of the *konda*-yam specifically. Badaga *seṇḍu* also means 'grass or straw ring,' Basque *bola* also 'skittle' and 'ball-bearing,' and Khalkha *bombyge(n)* also 'bomb' and 'pellet used for crossbows.' Similarly, Welsh *pêl* also means 'pellet' and 'pill.' Cheyenne *htóhtséme* ~ *hohtsemo* also denotes a 'netted hoop,' and Lake Miwok *polóolo* colexifies 'ball' with 'dry oak ball' as well as 'baseball' and 'Ball Dance' specifically. Hawaiian *pōpō* also denotes a 'round mass,' 'cluster,' or 'bunch' generally, *pōka'a* (containing *ka'a* 'to turn') also means 'coil,' 'roll,' and 'spool' inter alia. The basic meaning of Lenakel *nouanetpinami*- is 'urinary bladder,' which is reflected in the term's morphological structure: it is analyzable as /noua-netp-nami-/ 'fruit-belly-urine' (presumably inflated animal bladders are used as balls in games).

Otherwise, terms for 'ball' may contain verbs meaning 'to roll' or to 'roll around' (Efik, where the term also denotes the game played with the ball) or 'to kick,' as in Comanche *naʔsuhpeeʔ* 'small native-american football,' which is analyzable as /na-suh-peeʔ/ 'REFLX-kick-fall,' 'kick' is also the meaning Muna *sepa* 'ball plaited from rattan' assumes when used as a verb. By provenience contiguity, Rendille *imbíra* also denotes an object made of rubber or plastic generally, while Bora *máákiñiu* contains *máákiñ* 'rubber' and the classifier for small round objects *-iu*. In three languages of the New Guinea area, Buin, One, and Sko, there is a curious association between 'ball' and 'citrus': in Buin, 'ball' is colexified with a "kind of tree: a tree with spherical, lemon-like fruit about 10cm in diameter," One *malwa tala* ~ *maula rala* is analyzable as /malwa tala/ 'citrus round.thing,' and Sko *hangléúe*, which is a dvandva compound analyzable as /hang-lèue/ 'coconut-peanuts,' colexifies the meanings. Muna, spoken in the same broader region provides a hint towards an explanation: *golu lemo* is glossed as 'citrus fruit used as ball' (*golu* 'ball,' *lemo* 'citrus fruit'). Pawnee *rariickiriʔ* also denotes the 'baseball' and the associated game specifically which is also the case in Lake Miwok and Hawaiian, just like Buli *gbeli*, Carib *bal*, and Bwe Karen *bóló* also denote the 'football' specifically. Ngaanyatjarra *purrpurl(pa)*, borrowed from English *football*, also denotes the game of the same name as well as "football carnival," (*tjaputjapu* likewise denotes both 'ball' and the game of football). Dongolese Nubian and Japanese

colexify ‘ball’ with ‘bowl,’ in Japanese due to phonological collapse of the English source words *ball* and *bowl*.

Other patterns include: Efik *obön* is also glossed as ‘musketo,’ Berik colexifies ‘ball’ with ‘forefinger,’ Mali *vaidebunggi* is used to refer to something of tennis-ball size and is also the name of a tree species, while Rotokas colexifies ‘ball’ with ‘name’ and ‘letter.’ Greek *mpállá* also means ‘bale,’ and Japanese *kyū* also ‘nine.’ A literal rendering of Carrier *nekhek* would be ‘uses to be tossed about,’ and Comache *na?mahpe?e?* would be translated literally as ‘object thrown by hand.’ Kiowa colexifies ‘temple’ and Santiago Mexquititlan Otomí *lobo* ‘wolf’ (due to phonological collapse of Span. *globo* and *lobo*). Central Yup’ik (Nunivak Island dialect) *yuguaq* is analyzable as /yug-(ng)uaq/ ‘person-imitation,’ and San Lucas Quiaviní Zapotec *pelo’t* also means ‘ball game.’ Embera colexifies ‘ball’ with ‘whole’ and “surrounding district.” Miskito *dans pulanka kum* is analyzable as ‘dance play one.such.’ Imbabura Quechua *muyu-ndin* is analyzable as ‘seed-with,’ whereas Hani *siilpuq* might contain *puq* meaning ‘to swell’ inter alia (*siil* is ‘gold, yellow’). Rotuman *poro* also is the name of a bush inter alia, and the plural of Samoan *polo* also means ‘roller bearings.’ Finally, Vietnamese *bóng* also means ‘shadow’ inter alia.

#### 74. *The Bed*

Representation: 88%

Motivated: 46.8%

Thereof Analyzable: 28.7%

Thereof Colexifying: 17.7%

Thereof by Contiguity: 30.0%

Thereof by Similarity: 10.0%

Recurrent associated meanings: sleep, furniture, lie/lie down, place, mattress, tree/wood, mattress, ground/floor, (bed-)room, quilt, thing, spread/spread out, nest, hammock, dream

Commonly, words for ‘bed’ contain verbs meaning ‘to sleep’ or ‘to lie’ (for which see Buck 1949: 480 in Indo-European), and they may be either of the derived or lexical type. Terms derived from ‘to sleep’ are found in Muna, Cahuilla, Chickasaw, Nez Perce, Central Yup’ik (Nunivak Island dialect, otherwise “sleeping bag, bedroom”), Cashinahua, Cavineña, Ancash Quechua, Fijian, Rotuman, and Samoan (here, there is a more complex term: *ulu-moe-ga* ‘enter-sleep-NMLZ’). If terms are, however, of the lexical type, the second element may differ: it may simply be ‘thing,’ as in Mbum *fê-nâm* ‘thing-sleep,’ ‘place,’ as in Yoruba *ibùsùn*, which is analyzable as /ibi-isùn/ ‘place-sleep,’ as well as in Wichí (and note that Ngaanyatjarra, Abzakh Adyghe, Khalkha, and Fijian colexify ‘bed’ with ‘place,’ and Anggor inter alia with ‘place,’ ‘chair’ and ‘floor’), or it may be the name of some other piece of furniture, as in Dadibi *pibo sai* ‘sleep floor/table,’ and also in Kyaka, Rao, Toaripi, and, questionably, Japanese (it is in fact relatively frequent for terms for ‘bed’ to also denote another piece of furniture or furniture in general, this pattern is found in Noni, Anggor, Baruya, Burarra, Mali, One, Rotokas, Sahu, Badaga, Nuuchahnulth, Tuscarora, Lengua, Rama, Hawaiian, Rotuman, and Bislama). Otherwise, complex terms involving ‘tree’ or ‘wood’ alongside ‘sleep’ are also found: Buli has *gaduok* /goa-duok/ ‘sleep-wood,’ Kyaka *luu palenge isa* ‘sleep ?? tree/wood,’ Copainalá Zoque *nguy* /ɲu-cuy/ ‘sleep-wood/tree,’ Kaingang *ka*



*krē* ‘wood criciúma’ (criciúma is a type of Bamboo; the complex term also denotes a ‘field to beat beans’), and Maxakalí *mîp-xap*, analyzable either as ‘wood-stone/bead/seed’ or ‘wood-weave/knit/sew’ (note also Yir Yoront *yo-way* ‘high branches of a tree, stretcher, bed,’ which is analyzable as ‘tree-high.up’ and Itzaj *tus’bil-che’ ~ tus-che’* ‘wood pile, rack, bed,’ analyzable as ‘stack(ed)-wood’). Furthermore, the Bororo term *boe enu pa* contains both terms for ‘thing’ (*boe*) and ‘place’ (*ba*) alongside a verb meaning ‘to sleep’ (*nu*). In Ngaanyatjarra, Aymara, Bora, and Hawaiian, ‘(to) sleep, sleeping’ and ‘bed’ are directly colexified (in Hawaiian also other meanings are), in Aymara by the analyzable term *iki-ña* ‘dream-INTR.’ In fact, Hawaiian colexifies ‘bed’ also with ‘dream,’ and relevant terms in Bora, Cashinahua, and Toba may or may not do the same, since the Spanish gloss ‘sueño’ is ambiguous between ‘sleep’ and ‘dream.’ Semianalyzable terms involving a verbs meanings ‘sleep’ or ‘go to sleep’ is found in Nez Perce, and one featuring a noun ‘sleep’ in Toba.

The other major association, that with verbs meaning ‘to lie’ or ‘to lie down,’ betrays the same mixture of derived and lexical terms. The derived variety is found in Efik (where the term also denotes a ‘camp’ and a ‘rendezvous’), Carrier, Upper Chehalis, Ineseño Chumash, Kiliwa, Lakhota, Wintu (based on a verb meaning ‘lie on ground’ more specifically), Yuki, Jarawara, Fijian (where the derivation base is glossed as ‘to lie in a place’ more specifically, and hence the derived term also has the more general meaning ‘location, place, position’), Malagasy, and Rotuman. For instance, Malagasy has *fandriana*, analyzable as /fa-àndry-ana/ ‘NMLZ-laying.down-NMLZ.’ In Cheyenne, ‘bed’ is *šéšestôtse* /šéše-hestôtse/ ‘lie-thing,’ in Haida it is *tay daan* ‘lie place,’ colexifying ‘basking place’ and ‘rookery.’ In Tetun *tobafatin*, *fatin* is ‘place’ and *toba* can actually mean both ‘to lie down’ and ‘to sleep,’ and the same situation is encountered in Kwoma. Semianalyzable terms with ‘lie’/ ‘lie down’ occur in Biloxi and Lengua (where the identifiable root also means ‘to sit’). The Biloxi term in fact colexifies ‘bed’ with ‘mattress,’ and this pattern is also found in Buin, Abzakh Adyghe, Badaga, Upper Chehalis, Wintu, Tsafiki, and White Hmong, while Cubeo has *paraino*, with the root *parai-* also occurring in the term for ‘mattress’ and classifiers differentiating between the two referents. Similarly, Kiowa, Arabela, and Aymara colexify ‘bed’ and ‘quilt.’

Another recurrent association realized by complex terms is that with the meaning ‘to spread,’ ‘to spread out’ (a pattern also evidenced in Old Norse and Church Slavonic, Buck 1949: 480): Swahili *kitanda* contains the verb *tanda* ‘to spread, stretch’ and a noun class prefix, while Rotokas *urua* is derived from *uru* “spread something out to sleep or lay [sic] on,” Ineseño Chumash *suwaskinimu* is analyzable as /suwaskin-mu/ ‘to.spread.something.open-DERIV,’ and Badaga *ha:sike* contains *ha:si* ‘flat, to spread out’ (a similar pattern may be discernible etymologically in Koyraboro Senni). In contrast, an association exclusively realized by colexification is that with ‘(bed-)room,’ occurring in Ngaanyatjarra, Pawnee, Tuscarora and Guaraní, where the term in addition colexifies ‘nest,’ as is the case in Manange. Anggor, Baruya, Wintu, and Tsafiki colexify ‘bed’ with ‘ground’ or ‘floor’ (Wintu *tEh* also means ‘spread,’ ‘hay,’ and “hind legs of dead animals spread and dragging” and Lake Miwok *wéja* also ‘base, bottom’ inter alia; compare also Kashaya *cahti* ‘bed’ and *ca-* “be sitting on ground, floor”?), while Nuuchahnulth *čim?iŋ* is analyzable as /čim-iiŋ/ ‘right/ready-ON.THE.FLOOR/IN.THE.HOUSE,’ and Wintu *pominpanas topi*

contains *pominpana* ‘to lie on ground’ (for the association with ‘ground’ in Baltic, see Buck 1949: 480). Furthermore, Cashinahua and Piro colexify ‘bed’ with ‘hammock,’ and Dongolese Nubian with ‘bedclothing.’

Other associations include: Hausa *gado* inter alia also denotes the notions of ‘inheritance,’ ‘bargain,’ while Koyraboro Senni *daari* also denotes a ‘bed-roll’ and ‘bedding.’ Ngambay *tuwa* also denotes the ‘straw to make a mat with’ and other things, and Baruya *minyē* can also refer to the “lower side of an ambush.” Muna *koe* “wooden or iron bed” also means ‘do not,’ while Ngaanyatjarra *ngurra* can also mean ‘camp’ and ‘living area.’ Badaga *me:de* also denotes ‘cane, bamboo,’ Chukchi *ajkol* also ‘skins to sleep on,’ Khalkha *oru* inter alia also ‘vacant place, vacancy’ and ‘trace,’ and Welsh *gwely* also ‘family.’ Lake Miwok *káama* also means ‘crib,’ and Nuuchahnulth *čim?it* is also “Chimihl, the name for the passage between Congreve Island and the shoreline of Barkley Sound.” Santiago Mexquititlan Otomí *xifi* also means ‘to explain,’ and Pawnee *kusaahkus* also means “be the site of a former dwelling, camp ..., be a campsite” and “be a field, playing field.” This term is in turn based on the term *kusaar*, which is derived from *kus* ‘to be sitting, be living,’ and also denotes a ‘seat’ and ‘place’ generally as well as a station in a ceremony specifically. Tuscarora colexifies ‘bed’ with ‘position,’ ‘space,’ and ‘stage,’ and San Lucas Quiaviní Zapotec *ca'mm* also denotes an “iron cup filled with gunpowder and exploded like a firecracker.” Central Yup'ik *aci* ~ *aciq* also means ‘area below, area under,’ Abipón *n-aol-Ra* is analyzable as ‘POSS.INDEF/3SG-bring-ABSTR,’ and the Arabela term *maqueja* seems to be derived from *maque* ‘footprint.’ Cayapa colexifies ‘bed’ with ‘mosquito net,’ Chayahuita *pě'sara* is analyzable as /*pě'sa*-*ra*/ ‘palisade-CLASS.SMALL.THINGS.’ Miskito *krikri* may also refer to ‘bunk’ and ‘tapestry,’ while Sáliba *jahaxoode* contains *jaha* ‘feet.’ Toba colexifies ‘bed’ with ‘totora sedge raft,’ and Hani *hhaoqzao* contains *hhaoq* ‘pillow’ and *zao*, acting inter alia as the classifier for beds. Kapingamarangi *hada* also means “platform of outrigger boom,” while Bwe Karen *lo ð-mí= ð-a* is analyzable as ‘stone rest-when=rest-eat.’ Finally, Hawaiian *moe* inter alia also means ‘to lie in wait, ambush’ as well as ‘to marry, sleep with,’ Sedang *xóang* also means “to solve a problem, to divide,” and Tetun *kama* also ‘cradle.’

#### 75. The Belt

Representation: 86%

Motivated: 36.4%

Thereof Analyzable: 26.8%

Thereof Colexifying: 9.3%

Thereof by Contiguity: 22.9%

Thereof by Similarity: 0.3%

Recurrent associated meanings: tie/bind/wrap/fasten, waist, strap, leather/skin, circle/loop, belly/stomach, bandage, rope, thing, middle, body

Most commonly, terms for ‘belt’ (‘sash,’ ‘gird,’ ‘girdle,’ ‘waistband’) are derived from verbs meaning ‘to tie, bind, wrap, fasten,’ or contain an element meaning ‘waist,’ or both (note that in Indo-European, ‘belt’-terms are derived from the root *\*yōs-* ‘gird,’ which may have originally been *\*yōu-s-*, an extension of *\*yeu-* ‘bind,’ Buck 1949: 434). Pure derived terms, such as Efik *u-bōp* ‘NMLZ-tie’ occur not only in this language, but also in Khoekhoe, Ineseño Chumash, Tuscarora, Central Yup'ik, Cashinahua (derived from a verb meaning ‘tie around

waist' more specifically), Cubeo, and Toba. Muna and Hawaiian colexify nominal 'belt' with verbal 'tie (around waist),' and similarly, Bezhta has *tic'yo*, a past participle of a verb meaning 'to put on,' Tuscarora *uyɛʔkwihčreh* contains the verb *-yɛʔkwih-* 'to gird,' and semianalyzable terms featuring constituents meaning 'tie up' and 'tied' occur in Upper Chehalis, Copainalá Zoque, and Hawaiian. There are also complex terms with more than one lexical element, where one of them is a verb meaning 'to tie.' Kiowa has *tən-p'ɥ-ga* 'gap-be.tied-NOUN.POSTFIX,' Nuuchahnulth *tapwanim* contains *tap* 'tie about' and *im* 'thing' (where *tap* in fact by itself means 'to gird, belt, tie about'), Bororo has *boe ekajewu boe* 'thing' and *kajewu* 'bind,' and similarly, Kaingang *vēsogfin ja* is analyzable as 'bind.oneself thing.' Lengua has *aptete nipthlit*, with *tete* referring to "anything which is tied" and *ninthlit* to the 'waist,' Cubeo *jārióicāva* is analyzable as */jārióñũ-cāva/* 'tie-CLASS.CLOTH-RIBBON.LIKE' and *jārió-me* as */jārióñũ-me/* 'tie-CLASS.LIKE.THREAD,' and in Hawaiian, one of the words for 'belt' is *kama 'aha* 'tie string.' More frequent, however, is the combination with 'waist' or 'stomach,' as in Abzakh Adyghe *bğə-rə-pχ(e)* 'waist-INSTR-tie,' which is also found in Mbum (where an additional element 'cord' is present), Sora, Upper Chehalis (where 'waist' is colexified with 'middle,' compare Fijian *i vau ni tolo-na* 'DERIV tie POSS middle.part-POSS'); similar denominations are Malagasy *fehikibo*, analyzable as */fèhy-kibo/* 'tying/knot-stomach' and Tetun *futu-kabun* 'bind-belly.' Jarawara has *makari tosi* 'cloth/clothing waist,' and furthermore, Aguaruna and Yaqui feature a monomorphemic term colexifying 'waist' and 'belt.' Derived terms with the lexical basis meaning 'waist,' such as Abipón *aat-Reki* 'waist-LOC,' are also found in Basque and Aguaruna. Other complex terms where one of the constituents means 'waist' or 'stomach' occur in Buli (*chiak gbain* 'waist leather'), Mbum (*sàl tàk* 'cord waist'), Kolyma Yukaghir (*aṇdil-amdi*: 'waist-bedding'), Yaqui (*toma jisumiam*, analyzable as */toma jisumia-im/* 'belly package<sup>8</sup>-PL'), Guaraní (*ku'a-sâ* 'waist-rope' and *ku'a-kua-ha* 'waist-have.holes-AGT'), Miskito (*maisa-wila* 'waist-ribbon'), Ancash Quechua (*tsiqlla watu* 'waist thread/strap'), and Vietnamese (*dây lưng* 'cord waist'). Semianalyzable terms with 'waist' occur in Bororo, Cayapa, Guaraní, Huambisa, Sáliba, and Wayampi, and the Ineseño Chumash term *tiwošokuš* contains the verb *tiwošok-* 'to wrap around the waist' and is literally translated as "something wrapped around the waist;" Copainalá Zoque has a semianalyzable term where the identifiable constituent means 'belly.'

Alongside the Buli term mentioned above, the meaning 'leather' or 'skin' also figures in Bora *ócáji-múhe* 'cow/tapir-CL.skin/leather,' Rama *bípuk* /bip-uk/ 'cow-skin' as well as in Efik, where 'leather' is colexified with many artifacts made from leather, among them 'belt' (the relevant term also has other meanings). Ngaanyatjarra, Lesser Antillean Creole French, Itzaj, Wintu, Cayapa, Macaguán, Sáliba, Rotuman, Samoan, and Bislama colexify 'belt' with 'strap' (see Buck 1949: 434 for evidence from Romanian), and three languages of the sample, Khalkha, Carib, and Hawaiian, colexify 'belt' with 'circle' and/or 'loop' (among other meanings by a number of terms in Hawaiian). Tetun *faixa* is also used with the meanings 'bandage,' 'lane on highway,' and 'track of record,' and Samoan *fusi* also denotes a 'bundle' or 'bandage' as well as "championship, final." Meyah *márféb efagá* is analyzable as 'cord body,' and Arabela *cajiniocuaque* as */cajinio-cuaqueya/* 'middle-body.'

<sup>8</sup> Original gloss is 'package.'

Other associations include: Hausa *'damara* is also the name of a “geometrical figure consisting of two interlaced triangles” and, dialectally, “[a] coloured glass bangle.” Noni *kecaw* appears to be derived from *caw* ‘select/choose’ by the noun class prefix *ke-*, Dongolese Nubian *gāṣ* also means to ‘coquet, flirt, mince, simper, be spoilt’ (Armbruster 1960: 131 relates the senses by way of the “underlying common notion” ‘embrace’). Yoruba *òjá* also denotes a ‘head-tie,’ Angkor *titapuri* has the meaning ‘bracelet’ and might also be capable of referring to the ‘belt,’ and a literal translation of Dadibi *kibu wali* seems to be ‘pig wind.’ Muna *bhida* also means ‘shroud.’ Yir Yoront *maq* colexifies ‘belt’ with ‘bottom, lowest part,’ and *monporm* is the Yir Yoront term for both ‘possum, fur of possum’ as well as a belt made from possum fur. Japanese *obi* is a nominalization based on a verb meaning ‘to wear,’ and Welsh *gwregys* also may refer to a ‘truss,’ and *bad*, an obsolete term, also means ‘plague.’ Biloxi *a'xkido'ni* contains a verb meaning ‘to wrap,’ Cahuilla *tépaqal* is derived from a verb meaning ‘to tighten a belt,’ while in Comanche, *kohinehki?* ~ *kohineeki?* has been extended to also denote a ‘G-string.’ Kashaya *pha?šati?* appears to contain the instrumental prefix *pha-* ‘by wrapping’ and *?šat-* ‘to hit hard and hurt.’ Lake Miwok *mitúpponi* is analyzable as /mítu-pponi/ ‘count-AGT.’ Tuscarora *ye?nhēhθu?narhúhstha?* contains *-(i)?nhēhθ-* ‘rawhide strip’ and *-u'narhu-* ‘hook.’ Wappo *pá'ha?* also means ‘straw’ (presumably due to the collapse of Span. *paja* ‘straw’ with *faja* ‘strip, waistband’), whereas Wintu *lakum* contains *lak* “get caught, trip, hook; embrace, pinch,” and *la'q čopçi* is used for rattlesnake skin that is worn as a belt specifically (and is indeed analyzable as ‘rattlesnake skin’). Bororo *aie-wora* appears to be derived from *wora* ‘leg’ and also may refer to a ‘Bororo tanga.’ Lengua *yukma tama* is analyzable as ‘skin.skirt string,’ and Fijian *i oro* as ‘DERIV clasp.’ Hani *juqzaq* appears to contain *juq* ‘loose’ (also meaning ‘to throw’) and *zaq*, acting inter alia as a classifier for bundles. As a verb, Kapingamarangi *duu* means ‘to stand’ and ‘to stop,’ and Lenakel *katovit* is derived by the instrument nominalizer *k-* from the verb *atovit* “to put on clothing by wrapping it around self.” Rotuman *fāli*, as a verb, also means “to thresh with a belt or strap.” Bislama *strap* also denotes a ‘seat belt’ and ‘fan belt’ inter alia, and Yay features semianalyzable terms involving a constituent meaning ‘man.’

#### 6.2.2.76. *The Boat*

Representation: 70%

Motivated: 33.1%

Thereof Analyzable: 10.7%

Thereof Colexifying: 23.3%

Thereof by Contiguity: 7.1%

Thereof by Similarity: 6.9%

Recurrent associated meanings: canoe, water, trough, raft, vessel, collamon, airplane, vehicle, fire

Motivated terms for ‘boat’ (which is not distinguished here from ‘ship’) are relatively rare. Quite common are patterns of colexification that are somewhat similar to that of English *vessel*, that is, terms that denote both a receptacle and container (for goods or other things) and a means of transportation. This situation is encountered in Basque and Central Yup'ik (and in Ancient Greek and Sanskrit, Buck 1949: 730); the Basque term also denotes a ‘case,’ ‘sheath,’ ‘carton,’ or ‘pot,’ and the Yup'ik term also a ‘tray’ specifically. Further-

more, Hausa, Mbum, Khalkha, Chickasaw, Pipil, and Tuscarora colexify 'boat' with 'trough.' The relevant Hausa term *jirgi* is also unique in being also applicable to a railway train (the two relevant Hausa terms have still further meanings). Further, Khoekhoe *tgaub* is also the name of an 'elongated bowl,' Gurindji *kartiya* and Yir Yoront *pinarr* are also used as the name of the 'coolamon,' an indigenous Australian carrying vessel, as well as, in Yir Yoront, for a "deep wooden oval dish," Badaga *teppa* also denotes a "temple pond" as well as an "artificial tank," Wappo *khéye* also denotes a 'cradle basket,' and Embera also a 'chest,' while Fijian *waqa* inter alia can also refer to a 'box, case, container' generally.

Complex terms for 'boat' are of a variety of structural types, and most often make reference in some way to 'water.' Kanuri has *mààrá njî-bè* 'vehicle water-of,' Yoruba *òkò-oju-omi* 'vehicle-eye-water' (alongside *òkò kekere* 'vehicle little'), and 'boat' and 'vehicle' generally (as well as 'conveyance') are colexified in Quileute. Kaluli has *hɔ:n kɔ:su* 'water airplane' (this term denotes a modern-type boat introduced in the colonial era; note also that 'airplane' and 'boat' are colexified in Khalkha, which features a very general term for any kind of vehicle), Blackfoot *aahkioohsa'tsis* /*yaahkioohsi-a'tsis*/ 'travel.on.water/travel.by.boat-INSTR,' Cheyenne *amóhehstestôtse*, containing the prefix *am-* 'along by water' and *hestôtse* 'thing,' Chickasaw *okokaaittanowa*, analyzable as /*oka'-okaaittanohówa*-/' 'water-walk-NMLZ,' Comanche *pawobi*, analyzable as /*paa-wobi*/ 'water-board' (in the Kwahere dialect, there is an additional constituent meaning 'horse'), Santiago Mexquititlan Otomí *bojá dehe* 'iron/car water,' and Piro *gonu yapachro*, containing *gonu* 'water' and *ya* 'go.' There is a semianalyzable term in Guaraní. Moreover, there are two languages of South America with complex terms where one of the constituents is 'fire': Bora has *cújúwa-mi* 'fire-SCM.transport' and Wayampi *tata-l-ena* 'fire-of-place,' which colexifies 'fireplace.' In both cases, the terms denote a 'steamboat' specifically.

Santiago Mexquititlan Otomí has *motsa dehe* 'canoe water,' and generic terms for 'boat' are colexified with 'canoe' in Buli, Hausa, Mbum, Noni, Muna, Tasmanian (all varieties expect the North-Eastern one), Biloxi, Carrier, Cheyenne, Ineseño Chumash, Comanche, Haida, San Mateo del Mar Huave, Kiowa, Lakota, Nez Perce, Tuscarora, Yaqui, Abipón, Guaraní, Sáliba, Toba, Wayampi, and Fijian, and with 'raft' in Buli, Dadibi, Badaga, and Yuki.

Other associations include: Hausa *komi* also denotes the "beds of an irrigated farm," while Dongolese Nubian *kúb* is also used with the meaning 'shuttle in weaving.' Lavukaleve *fela'koe* also means 'village.' Muna *kapala*, a Loanword from Bahasa Indonesia, also means 'leader, chief,' and indigenously also 'to sit on something raised,' while *bhangka* is also the name of a constellation and means "inner part of belly." Abzakh Adyghe *q<sup>o2</sup>əhe* contains *he* 'to carry away.' Welsh *cwch* also means 'beehive.' Kiowa *kā'bout* contains *kā-* 'to swim, to go by boat,' and Oneida *kahuweyá* is also the term for the 'black ash.' Pawnee *rakuuhuuru* is analyzable as /*rak-huuhuuru-u*/' 'tree/wood-floating-NOM.' Chayahuita *panca nansha marë pa'tërin-so* is (semi-)analyzable as 'big ?? sea leave-3SG.SUB.' Toba *lllicta ~ lllicota ~ lllicota* is derived from *illigot ~ illogot* 'rows,' while Wayampi *ia* is also the name of an ant species. Fijian *velovel* is also the name of a string figure, Hani *loq* also means 'to rinse a container with water' and acts inter alia as a classifier for irrigated fields, Bwe Karen *khli* is

also the name of a month, Hawaiian *moku* also means ‘island’ inter alia, while Kapingamarangi *waga baalii* (containing *waga* ‘canoe’) also means ‘grasshopper.’

### 77. *The Car*

Representation: 66%

Motivated: 41.9%

Thereof Analyzable: 27.8%

Thereof Colexifying: 15.2%

Thereof by Contiguity: 18.5%

Thereof by Similarity: 10.2%

Recurrent associated meanings: vehicle, cart/carriage, machine, roll, self, run, thing, train, motor vehicle, canoe, sled, twist, fire, house, ride, land

A very interesting pattern in terms for ‘car’ is that some of them contain elements meaning ‘move’ and ‘self’ (which is in some of the relevant languages conveyed by a reflexive marker). This is ultimately the literal meaning of the Graeco-Roman hybrid compound *automobile*, the constituents of which have just this semantics. This denomination also recurs in other languages alongside Greek, where it is (at best) semianalyzable in the modern language, for instance in Carrier *tigerh-nekhēs* ‘proceed-by.itself,’ and also in Nez Perce, while Kashaya has the somewhat similar *yu?dul q<sup>h</sup>ayam?*, analyzable as /yu?dul <sup>h</sup>q<sup>h</sup>ay-am-?/, containing *yu?dul* ‘self’ and <sup>h</sup>q<sup>h</sup>ay ‘run,’ and the literal translation of Lakhota *iyéčhikiyqke*, according to the consulted source, is “it runs by itself” (for the element ‘run,’ compare also Ngambay *né kàyn ngoru* ‘thing run fast,’ Samoan *ta’a-vale* ‘run.freely-bad/of.no.use/ordinary,’ and Kiliwa *(?)wa?kw-s-?+hin* ‘(DN+)house=WH-IRR-DN+run’). Tehuelche has the similar term *wawere:nk* /waw-?ere:-n-k/ ‘sole-walk-NMLZ-MASC,’ which can also refer to a ‘bachelor’ (there are further variants of this term). The presence of this pattern in many languages of the Americas almost suggests calquing, although, obviously, this would presuppose knowledge of the meaning of the constituents. Whether indeed these terms were calqued under European influence (perhaps mediated by missionaries?), coined independently, or a mixture of both remains an open question. Further, somewhat similar is the Fijian term *qiqi toso* ‘roll/vehicle move.of.itself,’ and a term containing a verb meaning ‘to roll’ is also featured in Chickasaw (*itti’ chanaa palhki* ‘wood roll be.fast’), while the association with ‘rolling’ is by direct colexification in Hawaiian (among other meanings); compare also the Ngambay term with a constituent ‘speed’ mentioned above as well as Cheyenne *ameohe-hestôtse* ‘go.by.quickly-thing,’ Central Yup’ik *akag-cuun* ‘roll-device.for’ (this term colexifies ‘wheel’ and ‘axle;’ there are other dialectal variants), and Hupda *pəpád-teg* ‘roll-thing.’

Cheyenne also has the alternative term *am-âho’-hestôtse* ‘along-by.heat-thing,’ and, somewhat similarly, there are two languages in the sample, Bora and Wichí, in which relevant terms feature an element meaning ‘fire’ (*cúújúwa-mi* ‘fire-SCM.transport’ and *wej itoj* ‘end fire’ respectively). In two sampled languages, terms for ‘car’ make reference to the loud noise it produces: Mali has *araun’ga mētki*, analyzable as /araun-ka mēt=ki/ ‘sound-M.SG in=3F.SG’ (this term may also refer to a motorcycle), and Acoma *tárárágá*, analyzable as /tárərəka-/ ‘be.roaring-INTR.’ In Efik and Kiliwa, there is a metaphorical transfer from ‘house’ to ‘car:’ the relevant terms in these languages are *ufök enañ makara* ‘house cow

European' and  $(?)wa?kw-s-?+hin$  '(DN+)house=WH-IRR-DN+run,' as already mentioned. Tuscarora has  $u\theta r\acute{e}h\check{c}reh$ , containing the root  $-(i)\theta(e)r-$  'to ride,' and a similar term, containing the root  $-?sle-$  'drag, ride, drive, trick someone' is found in the related language Oneida. Central Yup'ik  $nuna-kuar-cuun$  is analyzable as 'land-over-device.associated.with,' and Piro  $\acute{t}\acute{s}ixiyapat\acute{s}ro$  contains  $\acute{t}\acute{s}ixi$  'earth, land' and  $ya$  'to go.' Guaraní  $mba'e-jere$  is analyzable as 'thing-twist' (there is a further semianalyzable term with  $mba'e$ ), while 'to twist' is inter alia colexified with 'car' in Hawaiian.

There are also associations with other vehicles: Swahili, Abzakh Adyghe, Basque, Japanese, Khalkha, Bora, Kaingang, Malagasy, and Samoan colexify 'car' with 'cart' or 'carriage' (Bora by the analyzable term  $\acute{i}jch\acute{i}-e-mi$  'come.ashore-belong.to-SCM.transport;' it cannot be excluded that there is an error in the source), Comanche, Pawnee and Wintu with 'train' (Pawnee and Wintu have borrowed English *car* or the plural form *cars*), Upper Chehalis colexifies 'canoe' (while Kwoma has  $gaba veyi$  'whiteman/ghost canoe'), and Kildin Saami, Ineseño Chumash, Kashaya, Lake Miwok, and Wappo colexify 'car' specifically with 'machine' generally (all have borrowed the respective terms from contact languages which in turn ultimately go back to Latin *machina*; Kashaya also has optional complex terms on the basis of this term). Moreover, in two sampled languages spoken at high latitude, an association with 'sled' is found, in Central Yup'ik by colexification, and in Ket by the analyzable term  $\acute{e}y suul$  'iron sled.' Moreover, Koyraboro Senni, Rendille, Swahili, Japanese, Ket, Khalkha, Oneida, Tuscarora, Hupda, Lenakel, White Hmong, and Bislama colexify 'car' with '(wheeled) vehicle' in general, and Koyraboro Senni and Ngaanyatjarra with 'motor vehicle' (otherwise, the presentation does not differentiate between 'car,' 'truck,' 'bus' etc.). Similarly, Koyraboro Senni  $moobil-ize$  is analyzable as 'vehicle-child,' Yoruba  $\acute{o}k\acute{o} \acute{a}y\acute{o}-k\acute{e}l\acute{e}$  as 'vehicle fanciful-manner,' and Fijian  $qiqi toso$ , as already mentioned above, as 'roll/vehicle move.of.itself.'

Other associations include: Buli  $logri$  also means 'to give way,' 'to avoid,' and other things, while Kyaka  $karo$  also means 'dirt' and 'grime' inter alia (the meaning 'car' is due to borrowing from Tok Pisin). Muna  $mintoro$ , an obsolete term, is related to  $ntoro$  'to turn, rotate.' Ngaanyatjarra  $yurltu$  also inter alia means 'empty' and 'hollow tree,' the common denominator of the meanings probably being that a car is "hollow" in the sense that it provides space for sitting in. Sahu  $'oto$  also means 'to cut,' Basque  $auto$  also 'edict, judicial decree' and 'mystery play, religious play,' and Khalkha colexifies 'rook in chess' inter alia. Welsh  $car$  also means 'trap.' Blackfoot  $\acute{i}it\acute{a}sap\acute{o}pao'p$  contains  $sap$  'inside' and  $opii$  'to sit.' Upper Chehalis  $x^wiy\acute{u}y\acute{a}qs$  is derived from  $x^wiy-$  'to cut off;' the term might be a loan translation from Chinook Jargon. Comanche  $na?bukuw\acute{a}a?$  is analyzable as  $/na-puku-waa/$  'REFLX-horse-horn.sound,' while Pawnee has  $kiriiraawis$   $/kiriir-raawis/$  'anus-smoke' and variants of this term. Santiago Mexquititlan Otomí colexifies 'car' with 'tool' and 'iron.' Wintu  $p^hulurumes$ , a term found in the Trinity County dialect, contains  $p^hu\acute{l}$  'to blow' (literal translation provided in the source is "puffing one"). Yuki  $lulqm\acute{a}l$  is possibly literally 'oil puller.' Mandarin colexifies 'car' with 'chariot,' White Hmong  $hov$  also means 'short' (in the sense of 'vehicle' it is a borrowing from Lao), and Bislama  $trak$  (presumably due to collapse of Engl. *truck* with *track*) also means "footprint, spoor, track" and is also the name of a wheeled children's toy.

78. *The Chair*

Representation: 84%

Motivated: 53.0%

Thereof Analyzable: 40.6%

Thereof Colexifying: 12.8%

Thereof by Contiguity: 38.5%

Thereof by Similarity: 10.3%

Recurrent associated meanings: sit/sit down, furniture, wood/tree, place, thing, but-  
tocks/bottom, saddle, throne, situation, floor, dwell

Clearly, the most frequent association as realized by morphologically complex terms is that with verbs meaning 'to sit' or 'to sit down' (as in Indo-European, Buck 1949: 482). The association is by derivation in Efik (*i-tie* 'NMLZ-sit,' this term inter alia also means 'situation,' which is also colexified in Kapingamarangi, or 'state' figuratively), Burarra (where 'to sit on' is colexified with 'put one's weight on, tread on, step on'), Gurindji, Mali (colexifying also 'meeting' and "sitting of parliament"), Chukchi, Sora, Blackfoot, Cahuilla, Upper Chehalis, Chickasaw (where the relevant term colexifies 'toilet'), Ineseño Chumash, Kashaya, Nez Perce, Nuuchahnulth, Yuki, Central Yup'ik, Abipón, Bora, Carib, Cashinahua, Guaraní, Miskito, Imbabura Quechua, Tehuelche, Toba, Yanomámi, Fijian, and Samoan, while Hawaiian directly colexifies verbal 'to sit' with 'chair' inter alia. When terms contain a second lexical element, this is often 'wood' or 'tree,' as in Baruya *namwaalyita* /*namwaalimo-yita*/ 'for-sitting-wood.' This pattern is also found in Ngambay, Kyaka (where 'sitting' is colexified with 'living'), Yir Yoront (here the term contains more constituents: *yo-penpn pam nhin+nh* 'wood-flat person/body sit.down+REL'), Ket, Kiowa, Lakota, and Copainalá Zoque (other terms involving a constituent meaning 'wood,' but not 'sit,' are Swahili's *kiti*, consisting of a noun class prefix and *mti* 'tree' and Itzaj's *k'an-che*, which is perhaps analyzable as 'support-wood.' Note also the similarity between Yoruba *àgà* 'chair' and *agà* 'tree'). Otherwise, 'thing' figures as the second constituent alongside 'sit' or 'sit down,' as in Katcha *nimo ma th-andane* 'thing GEN ??-sit,' also in Mbum, Ngambay, Dadibi, and Bororo ('thing' is also the meaning of the identifiable constituent of the Lenakel word for 'chair'). In still other languages, 'buttocks,' 'bottom' or the like is attested as the meanings of the second constituents, as in Cheyenne *táxe'ésééstôtse*, which is analyzable as /*táxe*-*'esé*-*e*-*hestôtse*/ 'upon-buttocks-sit-thing.' Similar terms are found in Pawnee and Bwe Karen (in addition, Toaripi has *kiri posa* 'buttocks platform,' and Tuscarora *uthečhráhkweh* contains roots meaning 'buttocks' and 'to collect'). Furthermore, Khoekhoe has *fnû-ai/nao-s* 'sit.down-front.bench-3SG.FEM,' Wintu *kenla-s-po'm* 'sit-??-land/ground/floor' (compare colexification of 'chair' and 'floor' in Angkor), Piro *tuplapiye* contains *tuplata* 'sit down, be seated' and *pi* 'rod,' Bwe Karen has *lo féná* 'stone sit,' and White Hmong *roojzaum* contains *rooj* 'article of furniture' and *zaum* 'sit.' Semianalyzable terms with 'sit' or 'sit down' are furthermore featured in Upper Chehalis, Central Yup'ik, and Lengua (where 'sit' and 'lie' are colexified).

In Embera, Hawaiian, and Samoan (in both Polynesian languages, 'to sit' is colexified with 'to dwell' inter alia), 'chair' is colexified with 'saddle' among other meanings (given the morphological structure, this also seems to be the meaning of Miskito *aras nila pila* 'horse back down'), while in Noni, Angkor, Basque, Khalkha, Nivkh, Nez Perce, the



Santo Domingo de Guzmán dialect of Pipil, Quileute, Tuscarora, Wappo, Aguaruna, Aymara, Bororo, Cashinahua, Embera, Huambisa, Hupda, Kaingang, Lengua, Piro, and Yanomámi, terms meaning ‘chair’ may also refer to any piece of furniture in general or denote another specific piece of furniture (other than ‘stool,’ which is disregarded here), such as a ‘bench’ or a ‘couch.’ Words for ‘chair’ also have the general meaning of ‘place (of something)’ in Efik, Angkor, Ineseño Chumash, and Bororo. Burarra and Aguaruna colexify ‘chair’ with ‘throne’ (compare the origin of ‘throne’ from an Ancient Greek word for ‘chair’ more generally, Buck 1949: 481).

Other associations include: Buli *zukupaglik* contains *zuk* ‘head’ and *kpagli* ‘to rest one’s head.’ The original meaning of the term was ‘headrest’ rather than ‘chair,’ a situation which is mirrored by synchronic colexification in Rendille. Mali *achut ngēthathengbēt* is derived from a verb meaning ‘to lean back.’ Sko has *fú-jéng* ‘post.of.house/corner-place’ (though also note *fú* ‘bottom of a four-legged animal’). Bezhta *q’ō* also means ‘anvil,’ and Welsh *cadair* also ‘cradle’ and ‘udder.’ The Biloxi term *ya’xoxo’ni* contains *xoxo* ‘to swing.’ Oneida *anitskwahlákhwa?* is analyzable as */an-itskw-hl-hkw-wa?/* ‘SRFLX-seat/part.of.body.one.sits.on-set.on.top.of/place.on-INSTR-??.’ Tuscarora *u?θkwéhseh* also denotes a “cutting block” as well as a “round block of wood, piece of a log.” Yaqui *banko* also denotes the ‘bank,’ and Aguaruna *ekeémtai* contains *ekeémi* ‘set on top’ and the instrument nominalizer *-tai*. Guaraní *apyka* contains *apy* ‘extreme point, deposit,’ and Huambisa *ekemtai* possibly *eken* ‘room, bedroom.’ Ancash Quechua *silla* (< Span. *silla*) may also refer to a ‘frame’ or ‘harness’ (original Spanish gloss is ‘montura’), Wichí *to-wej-w’et* is analyzable as ‘POSS.INDEF-queue-place,’ and Yanomámi colexifies ‘desk, console.’ Kapingamarangi *lohongo* also means ‘situation, status’ and ‘office.’ White Hmong *tog* also means ‘block,’ and Rotuman *nofo’a* also ‘chief.’ Sedang *táng* also means ‘to look for,’ White Hmong *tog* also ‘to sink’ and “half-way point,” Bislama *jea* (< Engl. *chair*) also means ‘ticket, seat in a plane’ as well as ‘to eat.’

#### 79. *The Clock*

Representation: 70%

Motivated: 60%

Thereof Analyzable: 27.6%                      Thereof Colexifying: 32.4%

Thereof by Contiguity: 51.4%                  Thereof by Similarity: 6.1%

Recurrent associated meanings: sun, hour, day, time, measure, bell, season/age, eye, imitate/imitation, watch, strike

There are a wealth of terms of different types for the ‘clock’ (or ‘watch,’ which is accepted as a proxy for this concept) that make reference in some way to the ‘sun’ (compare Latin *sōlārium*, derived from *sōl* ‘sun,’ Buck 1949: 1002). Burarra, Gurindji, Toaripi, Upper Chehalis (with a vowel change), Cheyenne, Comanche, Kashaya, Kiowa, Wappo, Wintu, Bora, Cubeo, Jarawara, Yanomámi, and Kapingamarangi directly colexify ‘sun’ with ‘clock’ (in some languages, as discussed in section 60, also with ‘moon’ and sometimes also ‘month,’ in Burarra, as discussed in section 57 and § 6.2.2.1., also with ‘star’ among other meanings, in Kiowa also with ‘summer,’ in Wappo also with ‘calendar,’ and in Jarawara also with

‘thunder’ and ‘lightning,’ see Dixon 2004: 71 for the history of this association). In New Guinea, associations with ‘eye’ are attested in Kaluli (*of-a:-si* ‘sun-GEN-eye’) and Meyah (*mówa eitéj* ‘sun eye’). In two languages of South America, Aguaruna and Wayampi, the words contain constituents meaning ‘sun’ and ‘to imitate’ or ‘imitation,’ while in Yuki and Cashinahua (where ‘sun’ is colexified with ‘hour,’ for other associations with ‘hour’ see below), the second element is a verb meaning ‘to watch.’

There are also other complex terms of the lexical type where one constituent is ‘sun.’ In Wichí, *tokafwala* is derived from *fwala* ‘sun, day,’ Baruya *nyihaanya* is literally ‘sun-go,’ Rotokas has *ravireo vetaveta-pa* ‘sun count-DERIV,’ Biloxi has *ina’do’hi’~ ina’do’ho’ni’*, containing *ina’* ‘sun’ and *do’* ‘to look at, see,’ in Carrier, *sa-dzi’* ‘sun-heart’ is ‘watch’ and *sa-dzi’-tco* ‘sun-heart-AUG’ is ‘clock,’ Chicakasaw has *hashi’ kanalli isht ithana-* ‘sun/moon move.to.a.new.location with know-NMLZ,’ Pawnee *sakuhka’iirus* is analyzable as /sakuur-kair-hus/ ‘be.a.day/sun-put.in/on-IPFV,’ Arabela has *pananu shanacutaja* ‘sun measure,’ and Huambisa *etsa nakumkamu* ‘sun drawing,’ Yanomámi has a redundant complex term involving *mothoka ~ motoka* ‘sun, clock’ and *mii* ‘see,’ the literal translation of which is “that which serves to see the sun.” The notion of measuring time, as in Arabela, also plays a role in the conceptualization of ‘clock’ in a number of other languages. In Japanese, ‘time’ is the other meaning figuring in complex terms of the lexical type: *to-kei* is analyzable as ‘time-measure.’ This is the only language with this particular configuration, though there are others which betray an association with ‘time,’ as is discussed further below. When ‘day’ rather than ‘sun’ is the meaning of the contiguous constituent, it is more frequent to have terms with the meaning ‘count, measure’ as the second constituent, as in Hupda *wág tã?kéy* ‘day measure.NMLZ’ (thus precisely parallel to Old English *dæg-mæl*, Buck 1949: 1003), and also in Upper Chehalis, Lake Miwok, Pawnee, and Quileute, where an additional element meaning ‘thing’ is present. Furthermore, Kiliwa has *maat=k-w-p-c-?+wir-u?*, which is analyzable as ‘REFLX=WH-MP-INST/MOUTH-DN+measure-OBL,’ and an easier palpable literal translation offered by the lexicographer is “it measures itself.” Other terms in which ‘day’ figures are Guaraní *arairû ~ arirû /ára-irû/* ‘time/day/sky-companion’ and Malagasy *famantaranandro*, analyzable as /fa-fànatra-ana-àndro/ ‘NMLZ-known-NMLZ-day.’ In Toaripi and Comanche, ‘clock’ is colexified with ‘day’ (and also with ‘sun,’ compare section 60), and Blackfoot has a term derived from a verb meaning ‘be day.’ Yoruba, Burarra, Khalkha, Hani, Samoan, Bislama, and Sedang colexify ‘clock’ with ‘time’ or ‘time of day’ specifically (Muna *dhamu* also means ‘be time for something’ when used as a verb as well as “tonic made of medicinal herbs”) and an analyzable terms with ‘time’ is featured, alongside Japanese as already discussed above, in Piro (*hohi himata-tšaro* ‘time know-??’); moreover, there are semianalyzable terms in Khoekhoe and Haida, with the other element diachronically related to a verb meaning ‘to see.’

Frequently, ‘clock’ is colexified with ‘hour’ (see Buck 1949: 1002 for Indo-European evidence, e.g. from Middle High German). This occurs in Buli, Hausa (also with ‘good luck’ inter alia), Dongolese Nubian, Swahili, Berik, Muna, Abzakh Adyghe, Bezhta, Khalkha, Laz, Kildin Saami, Santiago Mexquitlan Otomí, Central Yup’ik, Chayahuita, Tehuelche, Hani, Bwe Karen, Samoan, and Sedang (colexifying also other meanings). In addition, Basque has *ordu-lari* ‘hour-AGT,’ Ket *časaj* ‘watch’ is the plural form of *čas* ‘hour’

(as with Russian *čas*), Oneida *kahwista?ékta?* is analyzable as /ka-hwista?ek-ht-ha?/ 'NEUT.AGENT-be.the.time.or.hour-CAUS-HAB.' The Hausa and Khalkha terms for 'clock' have very broad semantic latitude, and may refer also to temporal concepts such as 'season' and 'age.' In four languages of the Old World, Yoruba, Khalkha, Sora, and Mandarin, 'clock' is colexified with 'bell' (common in Celtic and from this source, Germanic, but also Latvian, Buck 1949: 1003-1004); Sora also colexifies 'gong.' In Efik, the word 'clock' contains *mi'a* 'to strike' and reflexive markers, and similarly, in Tuscarora, *kawenét?ehs* contains *-wenét* 'iron' and *-(i)?e(k)-* 'to strike.'

Other associations are: Buli *bang* also means 'bracelet, wristlet,' and Hausa *sa'a* inter alia also 'good luck' and "a propitious time." Nez Perce *liklíríes* is analyzable as /liklírí-nes/ 'go.around-INSTR.' Toba *lhuaxashi* contains *hua* 'forearm, hand,' Tehuelche colexifies 'clock' with 'alarm clock,' Fijian *kaloko* also denotes a 'very large kava, or roll of sinnet' (the meaning 'clock, watch' is due to borrowing from English). Kapingamarangi *laa* also means 'sail,' Mandarin *biao3* is elliptical for *shou3-biao3* 'hand-meter,' and is hence also used for other measuring instruments. Samoan *uati*, as a verb, also means 'to watch someone,' and Bislama *klok*, rarely, also may refer to the 'flamboyant, flame tree,' "because its leaves close up at night."

#### 80. *The Glasses*

Representation: 58%

Motivated: 62.4%

Thereof Analyzable: 55.8%

Thereof Colexifying: 6.6%

Thereof by Contiguity: 31.0%

Thereof by Similarity: 24.6%

Recurrent associated meanings: eye, glass, mirror, see/look, metal, cover, imitation, put/set

Motivated terms for 'glasses' are, perhaps unsurprisingly, clearly more often analyzable than colexifying, and, even less surprisingly, these terms with very few exceptions have one constituent meaning 'eye' by contiguity. As for the second constituent 'glass' is common, as in Muna *mata tonde* 'eye glass.' Such terms are also found in Efik, Yoruba (where an additional constituent meaning 'vision' is present), Kyaka, Ngaanyatjarra, Yir Yoront, Cubeo (additionally suffixed with the classifier *-ru* for roundish three-dimensional objects), Miskito, Hawaiian, and Yay. Furthermore, Hausa, Sedang, and Vietnamese directly colexify 'glasses' with 'glass,' Welsh has *gwydr-au* 'glass-PL,' and White Hmong *tsom-iav* 'look-glass.' Also common are terms based on a metaphorical comparison with a particular object made of glass, namely the 'mirror.' This is the case in Yoruba, Kyaka, Ngaanyatjarra, Hawaiian, Yay, so that these languages fall both in the categories exhibiting an association with 'glass' as well as 'mirror.' Languages in which complex terms for 'glasses' are found involving constituents meaning 'eye' and 'mirror,' without colexification of 'glass' and 'mirror,' are Buli, Koyraboro Senni, Mbum, Abzakh Adyghe, Badaga (where the 'mirror'-word also means 'lens'), Santiago Mexquititlan Otomí, Mandarin, and Rotuman. In the latter language, for instance, 'glasses' are called *maf tiro* 'eye mirror.' Further, Hausa, Bislama, and White Hmong directly colexifies 'glasses' and 'mirror' (and Hausa, as noted

above, also colexifies the substance ‘glass,’ while the Bislama term colexifies ‘glasses’ and ‘mirror’ and may also refer to other objects entirely or partially made of glass). Things get somewhat more complicated when one notes that some languages, in particular ones in which ‘glass’ is an item of acculturation, colexify ‘glass’ with ‘metal.’ Yir Yoront and Cubeo are of this type, and therefore, relevant complex terms for ‘glasses’ also betray a lexico-semantic association with ‘metal.’ There are, however, also complex terms for ‘glasses’ with this pattern where ‘glass’ and ‘metal’ are not colexified. These are Japanese, Cheyenne, and Nez Perce, which, for instance, has *kicúynim sílu* /kicúy-nim sílu/ ‘metal-poss eye;’ in Cheyenne, there is an additional constituent meaning ‘thing.’ In Acoma and Central Yup’ik, glasses are conceived of as an ‘imitation’ of the eye (Yup’ik has a dedicated postbase with this meaning). Here, ‘glasses’ are called *?úwána’azáhi*, underlyingly /húwana’ani-záhi/ ‘eye-imitation’ and *iinguak ~ iiguak ~ iiguak /ii-(ng)uaq/* ‘eye-imitation’ respectively. In two languages of the Americas, terms make reference to the fact that one ‘puts’ or ‘sets’ glasses onto the eye: Kashaya *hu?uy dute?ti?* contains *hu?uy* ‘eye’ and *dute?* ‘to put,’ and in Kiliwa, ‘glasses’ are *ny-?+yuw-l=t+papu-u?* ‘POSS-DN+eye-ILL=OBJ+set-OBL,’ literally, according to the lexicographer, “things one sets on one’s eyes.” Terms which directly make reference to the fact that glasses are used to aid seeing are surprisingly few. Chickasaw has *ishkin ishpiša* ~ *ishkinshpiša* which is analyzable as /ishkin isht piša-/ ‘eye with see-NMLZ;’ a similar term is only found in Upper Chehalis and Wintu (where there is a further term containing an element ‘eye’ and ‘to catch fish in a net, hold out a net to catch fish’). Hausa features a derived term from a verb meaning ‘to look,’ Fijian has *mata-ilo-ilo* ‘eye-look.at.reflection-RED,’ White Hmong, as already mentioned, has *tsom-iav* ‘look-glass,’ and Rotokas *osireipava sisiro* appears to contain *osireito* ‘eye’ and *sisiro* ‘inspect, stare, look intently.’ Pawnee, Kaingang, and Yanomámi feature terms for ‘glasses’ involving constituents meaning ‘eye’ and ‘to cover,’ Pawnee, for instance, has *kiriktahkuuku?u?*, analyzable as /kirik-raarkuuku-u?/ ‘eye-covering-NOM.’ A very similar term is found in Yanomámi, and Kaingang has *kanē kri táv* ‘eye above cover.’ Other complex terms with one constituent meaning ‘eye’ are Noni *ε-jise bala* ‘6-eye foreign,’ Angkor *hoe himboari* ‘water eye,’ Baruya *kwaari’matinna* /kwaari’mata-tinna/ ‘plastic-eye,’ Ngaanyatjarra *kurungkatja*, which contains *kuru* ‘eye’ and *katja* ‘son’ and also means ‘eye ointment,’ Bezhta *hăydă*, which is grammatically the plural of *hăy* ‘eye’ (and hence, can also refer to the ‘eyes’), Nivkh *njaḡ-aḡs* ‘eye-gold’ (the term is used to refer to a small piece of metal which is put on the eye of the deceased, and has presumably been extended to cover ‘glasses’ later from there on), Kolyma Yukaghir *šöjd-aḡd’ə* ‘stone-eyes’ and *aḡd’əd-aḡbi*: ‘eye-shadow,’ Comanche *pui tsa?nika?* ‘eye underwear’ (but compare *tsa?atsitū* ‘to inspect?’), Bora *hálluji* /hálluu-ji/ ‘eye-CL.disc,’ Guaraní *tesa-joa(py)* ‘eye-together’ and *tesa-irû* ‘eye-companion,’ Hupda *kəwəg-tú?* ‘eye-immersed.NMLZ,’ Piro *sutsa-yhalu* ‘clutch/pinch/grip-eye,’ Wayampi *εapalitō* /εa-palitu/ ‘eye-sparkling,’ Wichí *tot-telhu-hi-s*, containing *telhu* ‘eyes’ and the locative suffix *-hi* ‘in,’ and Malagasy *solomàso*, analyzable as /sòlo-màso/ ‘substitute-eye.’ Finally, Upper Chehalis, Tuscara, Toba, and Samoan directly colexify ‘eye’ with ‘glasses,’ and semianalyzable terms are found in Kemtuik, Cahuilla, Piro, Yanomámi, and Hani.

Other associations are few: Hausa *madubi* is also used to express the affections of parents to their child as well as to refer to a sorcerer. Ngaanyatjarra *winta* (< Engl *window*)

also means 'window.' Carrier *naḱētšelya* is the plural form of *naḱētšel* 'monocle,' and Tetun *ókulu* also means 'binoculars' and 'telescope.'

### 81. *The House*

Representation: 97%

Motivated: 29.7%

Thereof Analyzable: 3.0%

Thereof Colexifying: 26.4%

Thereof by Contiguity: 13.0%

Thereof by Similarity: 7.7%

Recurrent associated meanings: village, nest, room, family/lineage, roof, household, shelter, place, receptable, company, canvas cover, tepee, tepee cover, post, land

The most frequent association for 'house' (or 'building' generally, often also denoting 'home,' an association which is ignored here, just like cases when terms denote a specific type of house) is, by configurational contiguity, that with 'village' or more generally a group of houses, occurring exclusively by colexification in Katcha, Ngambay (which colexifies also 'country'), Kwoma, Mali, Ngaanyatjarra, Yir Yoront, Badaga (colexifying "isolated settlement" more precisely, as well as 'young'), Nuuchahnulth, the Cuisnahuat dialect of Pipil, Cashinahua, Jarawara, Miskito, Wayampi, and Bwe Karen (compare the cognacy of Ancient Greek *οἶκος*, *οἰκία* 'house' with Latin *vīcus* 'group of houses, village' and other evidence from Indo-European reported in Buck 1949: 458). By meronymy, some languages use the same word for 'house' and 'room.' These are Hausa, Badaga, Bezhta, Comanche, San Lucas Quiaviní Zapotec, Aymara, and Bwe Karen. Also by meronymy, Burarra, Arabela, Cubeo, Jarawara, and Tsafiki colexify 'house' with 'roof' (this pattern is common in a group of Indo-European terms, Buck 1949: 458; Burarra also colexifies 'lid'). Presumably, this is also the motivation for colexification of 'house' with 'post' in Yanomámi (compare also Miskito *playa bila* 'post space,' which quite literally also denotes the 'space between two posts'). By functional similarity, Hausa, Khoekhoe, Kyaka, Muna, Badaga, Wintu, and Lenakel colexify 'house' with 'nest' (Muna also with 'web,' the Khoekhoe term *oms* is derived from the verb *om-* 'to build, construct,' compare Proto-Indo-European *\*domo-* ~ *\*domu-*, from *\*dem* 'build,' Buck 1949: 458). Nunggubuyu, Waris, Basque, and Haida colexify 'shelter' (Nunggubuyu "stringybark shelter or dwelling" specifically), and Pawnee and Tehuelche 'canvas cover.' Ket *ĩḡcus* contains *qu ʒ* 'tent,' Wintu colexifies these meanings (compare Slavic evidence reported in Buck 1949: 459), and Nez Perce *cóqoy* also means 'teepee top, smoke hole' ('teepee cover' is also the meaning of the relevant Pawnee term, and Comanche and Kiowa colexifies 'house' with 'teepee' directly; compare the cognacy of some Indo-European terms for 'house' with 'hide,' which go back to a root meaning 'to cover,' Buck 1949: 458). A pattern apparently particularly common in the Old World is metaphorical extension to 'family' or 'lineage,' found in the sample in Buli, Rendille, Abzakh Adyghe, Badaga, Basque, and Wintu, while Nuuchahnulth *maʔas* similarly also means 'tribe' (compare further the cognacy of the Ancient Greek and Latin terms mentioned above with Old Persian *viθ-* 'royal court, palace, family,' Buck 1949: 458). Similarly, Rendille, Kwoma, Badaga, and Lesser Antillean Creole French colexify 'house' with 'household,' and Central Yup'ik *enaʔ*, Bororo *eda* ~ *jeta*, Jarawara *tabori*/*taboro*, Miskito *watla*,

and Wayampi *ena* are also used with the general meaning ‘place’ (the Jarawara and Wayampi terms also mean ‘land’), and Badaga *mane* also with the meaning ‘ground.’ Basque and Greek colexify ‘house’ with ‘company, firm’ (the Haida term is glossed as ‘establishment,’ but it is unclear whether this actually refers to an enterprise). Lakhota *thí* also means ‘to live, dwell’ in verbal usage, while in Wintu, *bo’s* is derived from *bOh* ‘live, reside, remain, keep; stay, be in a sitting position, sit, dwell, stay’ by means of the generic aspect suffix *-s*. *Bo’s* also means ‘afterbirth’ and ‘navel’ inter alia. Relevant Dongolese Nubian and Khalkha terms colexify ‘house’ with ‘receptacle’ generally.

Other associations include: Efik *u’fök* is derived from *fök* ‘to spread a covering over, to cover, disguise,’ Hausa *’daki* is inter alia also used as a term for a year when counting the age of a horse, and Swahili *nyumba* contains *umba* ‘to create.’ Kwoma *aka* also means ‘cave,’ and *akama* also ‘social role’ and ‘home region.’ Kyaka *anda* also denotes an ‘open valley area’ as well as ‘adobe’ inter alia, Rotokas colexifies ‘cabin,’ Kosarek Yale *ae* also means ‘region,’ Yir Yoront *ngolt* also ‘wall’ of a house as seen from inside, Basque *etxe* also “lodging, shelter” and “frame, body” in the technical sense, and Nivkh *tyf* also means ‘quarters.’ Acoma *káçə*, when interpreted verbally, means ‘it is tall,’ Blackfoot colexifies ‘house’ with ‘lodge,’ and Ineseño Chumash *ma’m*, a rare word for ‘house,’ is also used adpositionally with the meaning ‘inside of.’ Lesser Antillean Creole French *kai* also means ‘fishscales’ inter alia, Tuscarora *unéhseh* also is used with the meanings ‘cage’ and ‘umbrella,’ Yana *-sza-* also means ‘upward,’ and Hupda *mǎy* also ‘comb, brush.’ Bororo colexifies ‘house’ with ‘palm leaf’ (similarly, Jarawara *yobe* is also the name of a palm species, and since Jarawara colexifies ‘house’ with ‘roof, thatch’ this may be the chain of associations that is also responsible for the association in Bororo), Embera with ‘inn,’ and Hupda with ‘burrow’ and ‘brush.’ The Maxakalí term *mīp-tut* is analyzable as ‘wood-mother’ or ‘wood-woven.net’ (compare Cashinahua *jive* ‘house, village, community’ and *ji* ‘tree, wood’). Wayampi *-əka* as a verb means ‘to cut with axe,’ and *(l)etā* also means ‘site of waterloving creatures’ (“Gîte des monstres aquaphiles”). Hani *laqhyul* contains *hyul*, meaning ‘inside, domestic’ (alongside ‘to be extremely comfortable’), Lenakel *nimwa* also denotes the ‘placenta, afterbirth,’ a ‘cocoon,’ and a ‘handle,’ while Hawaiian *hale* also means ‘institution’ as well as ‘host, hospitable person.’ Finally, Bislama *haos* also denotes the ‘bridge of a ship’ or the ‘cabin,’ as of a truck, and Sedang colexifies ‘house’ and ‘rainy season.’

## 82. The Key

Representation: 68%

Motivated: 37.0%

Thereof Analyzable: 33.0%

Thereof Colexifying: 5.3%

Thereof by Contiguity: 33.5%

Thereof by Similarity: 3.3%

Recurrent associated meanings: open/close, lock/keyhole, lock/unlock, door,  
house, tap

A common association for the ‘key,’ which is predominantly realized by morphologically complex expressions, is that with terms meaning either ‘open’ or ‘close.’ It comes in a variety of structural guises. Terms may be of the derived kind, as in Chukchi *ine-nwentet-*

*icyan* 'ANTIPASS-open-INSTR,' occurring also in Khoekhoe, Hausa, Swahili (by prefixation of a noun class marker), Burarra, Welsh (where the relevant term also means 'opening' and 'act of opening'), Blackfoot, Chickasaw, Central Yup'ik (Nunivak island dialect), Cashinahua, Guaraní, Piro, and Fijian (straightforward evidence for this patterns from Indo-European is only found in Celtic, Buck 1949: 469). Terms of the lexical type include Yoruba *ì-ṣí-kà* 'NMLZ-open-thing' (this term is not usual), Biloxi *tí' í'ṭpa'xo'ni' /ti í'-dupaxi'-o'-ni' /* 'house INSTR-open.door-make-CAUS,' and the analogous Kiliwa and Arabela terms *wa?h+kap-u?* 'house=3+open-OBL' and *tiootiu riatataja* 'door opener.' In Comanche, there is an additional constituent making reference to a pointed object present, Maxakalí has *pipkup mōhām-yīnnīn-'ax* 'metal.object ??-shut-can,' and Tehuelche *golk'o kotenwe*, *gonke:renwe*, *gonomk'enwe*, and *go:nko' kašomk'enwe*, all of which feature a term for 'door,' *gonk'o* ~ *go:nk'o*, and the instrument nominalizer *-we* alongside verbs meaning 'loosen,' 'shut,' 'open' and 'see' (a semianalyzable term with 'loosen' is also featured in Nez Perce). Hawaiian, due to the flexibility of its lexemes with respect to the syntactic slot they may occupy, colexifies 'key' with 'to open' inter alia. Further, Oneida has a term for 'key,' *aten-hotukwátha?*, derived from a complex verb meaning 'to shut or close a door' specifically. Semianalyzable terms involving either 'open' or 'close, shut' occur in Khoekhoe and Cheyenne, and note also the similarity between Laz *nkula* 'key' and *nkol* 'close.' A related pattern is that when constituents mean 'lock' or 'unlock' rather than the more general 'open' and 'close,' in fact, Rotokas combines all those meanings in its term for 'key': *tupa karu-pa* 'close/lock open/unlock-DERIV.' Efik features a derived term (*u-kpāhāre* 'NMLZ-unlock'), so do Khoekhoe (in which 'open' is also colexified with 'unlock'), Haida, and Piro; in Pawnee and Yuki, the term is of the lexical type (*rakcaakarikuku /rak-caakarikuk-hus/* 'tree/wood-unlock-IPFV' and *piṭ piṭ-ul* 'door lock-INSTR'). Furthermore, Upper Chehalis has *s-ḡólḡ=iyq* 'CONTINUATIVE-lock.up=house.' Semianalyzable terms involving a verb meaning 'to lock' are found in Kildin Saami, Hani (the other element means 'similar, true to life' when occurring on its own), and Yay. In a number of languages, 'key' is colexified with its counterpart, the 'lock' or 'keyhole.' This is the case in Hausa (where the relevant term is also dialectally used as the title of the principle advisor of the Emir inter alia), Nez Perce (by the term *wecé?ke?ts*, containing *we-* 'hit' and *cé?* 'be trapped, be blocked in, be stranded, be struck, be immovable,' colexifying 'hammer' additionally), Wappo, Wintu, dialectally in Central Yup'ik, and in Abipón. Similarly, Sko *long* also means 'hole, cave.'

As the discussion of terms with 'open' and 'close, shut' has shown, in many languages 'door' provides an additional contiguity anchor. Other terms with a constituent meaning 'door' are Abipón *l-aham-kate* 'POSS.INDEF/3SG-door-INSTR,' Bororo *baiporo epa* 'door instrument,' and Toba *lemaqte ñi lasom*, containing *emec* 'spoon' and *lasom* 'door.' In addition, complex terms involving a constituent meaning 'lock' are found in Miskito (*ki mita* 'lock hand'), Hani (*zovqdul*; *zovq* is 'lock' and *dul* can refer to a variety of longish objects, among them 'thigh' and 'pen'), Malagasy (*fanalahidy*, analyzable as */fanàla-hidy/* 'instrument.to.take.out-lock'), as well as Yoruba (*omọ àgádágodo* 'child padlock') and Manange (*Itantsa-Iama* 'lock mother'). Extension of kinship semantics to the pair 'key' and 'lock' is also found in Latvian (Buck 1949: 469), compare also Matisoff (1992), who shows that it is widespread in Southeast Asia.

Furthermore, a Ngaanyatjarra avoidance register term alongside other meanings colexifies ‘key’ with ‘knife,’ and Basque and Santiago Mexquititlan Otomí colexify ‘key’ with ‘tap’ among other meanings (the latter language also has an optional complex term with *bojā* ‘iron’).

Other associations include: Muna *kunsi* also means ‘button,’ “have a commercial partnership,” and ‘plot, to conspire.’ The Abzakh Adyghe term *ʔəncʰəbze* is analyzable as /ʔə<n>čʰə-bze/ ‘abstract<RELAT>-strap.’ Japanese *kagi* also means ‘hook’ (though the respective readings are distinguished in writing). Khalkha *tylki-gyr* is analyzable as ‘push-INSTR’ and also may refer to any “long-handled implement used to push snow, manure, etc.” Nez Perce *waholkaʔs* contains *we-* ‘with an implement’ and *hol* ‘to loosen,’ and Nuuchahnulth *ʔiqʷak* is analyzable as /ʔiqʷ-ʔak/ ‘untie-tool.’ Tuscarora *uhsəwáʔreh* also means ‘fork,’ ‘nail’ (on this origin for words for ‘key’ in Indo-European see Buck 1949: 468) and ‘needle,’ and Copainalá Zoque *wiʔtoquiuy* contains *wiʔtu* ‘to turn, be locked with key.’ Huambisa *yawi* is also used to refer to ‘safety pins,’ whereas Miskito *warbaika*, analyzable as /warb-aia-ka/ ‘turn-INF-DERIV’ also means ‘tongs.’ Fijian *kī* also denotes a fish species and means ‘to do’ (the meaning ‘key’ is due to borrowing from English), Malagasy colexifies ‘needle of pine tree,’ Rotuman *kī* also means ‘sepia’ inter alia, while Sedang *khúang* also means ‘drill’ and ‘brace.’ Bislama *ki* (< Engl. *key*) also means ‘gear’ and ‘spanner,’ as well as ‘key’ in the music-related sense and denotes the ‘pegheads’ of a guitar.

### 83. *The Knife*

Representation: 94%

Motivated: 20.9%

Thereof Analyzable: 18.7%

Thereof Colexifying: 2.5%

Thereof by Contiguity: 14.6%

Thereof by Similarity: 4.1%

Recurrent associated meanings: machete, cutting implement, cut, iron/metal/steel, sword, bamboo, razor, to skin, fish

As with other artifacts, terms derived from a verb denoting the principal activity that can be performed with that artifact, in this case ‘to cut,’ are also frequent for the meaning ‘knife’ (see also Buck 1949: 558 for Indo-European), but clearly not as frequent as with other artifacts, presumably due to the fact that knives are frequently indigenous tools and were present before the era of colonization. Derived terms are nevertheless found in Chickasaw, Kashaya, Abipón (where ‘to cut’ is colexified with ‘to shine’), Bora, and Ancash Quechua. Fijian has a derived term from a verb meaning ‘to cut with knife’ specifically, Khoekhoe has an obsolete term of this kind, and the association is recoverable etymologically for Dongolese Nubian, Kolyma Yukaghir, and Kiowa. Furthermore, San Mateo del Mar *nicojchay onij* contains *acooch* ‘cut’ and *onij* ‘meat’ (the language also has another term, *nitajcüy cüet* containing *ataag* ‘disembowel’ and *cüet* ‘fish,’ for which compare Yir Yoront *ngartyann* containing *ngart* ‘fish’ and *ye* ‘cut, slice’). In contrast, Toba *lpetegaxanaxat* contains the verb *petec*, meaning ‘to cut hair’ specifically (this term colexifies both ‘scissors’ and ‘razor,’ which latter association is shared by Nez Perce and Tuscarora). The meanings ‘knife’ and ‘cut’ are colexified in Bwe Karen and Samoan (here, by a polite term), and syn-



chronically semianalyzable terms of this kind are found in Kolyma Yukaghir, Kaingang, and Wayampi. Badaga, Sora, Cahuilla, Lake Miwok, Nez Perce, and Mandarin colexify 'knife' with 'sword.' Colexification with other cutting implements (such as a 'sickle,' a 'dagger,' or an 'axe'), or general terms for any sort of bladed tool is also found in other languages, namely Badaga, Khalkha, Lake Miwok, Nez Perce, Oneida, Rama, Toba, and Samoan. More specifically, Meyah, Sko, Basque, Cubeo, Hupda, and Jarawara colexify 'machete,' and four languages of South America also have complex terms on the basis of 'machete': Bora *n̄ĩtsúwá-wu* 'machete-DIM' (in fact, the word for 'machete' is derived from a verb meaning 'to cut'), Cavineña *cuchiro caca* 'machete small,' Chayahuita *cosëra'hua* ~ *cosora'hua*, probably containing *cosoro* 'machete for cutting grass' and a classifier suffix, and Yanomámi *sipara si* 'machete cover.' Six sampled languages, Baruya, Cahuilla, Kiliwa, Central Yup'ik, Kaingang, and Lengua colexify 'knife' with the material 'iron' or 'metal' and/or 'steel' (Baruya also with 'plastic'), while Cubeo has *tāu-ve* 'glass/metal-CLASS.SLIM.SLENDER.FLAT.OBJECT' and Toba *laicaua lạyi* 'metal/iron edge.' Four languages of broader Oceania, Kwoma, Lavukaleve, Lenakel, and Samoan (by a polite term) colexify, by provenience contiguity, 'knife' with 'bamboo' (Lenakel also with 'backbone' inter alia, and Samoan also with 'fishing rod'). In Kiliwa, the word for 'knife' is derived from a verb meaning 'to skin' (*na(y)=c+ruuw* 'child/small=INST/mouth+to.skin'), and Hawaiian colexifies these meanings alongside 'flint' and "to stand on edge."

Other associations include: Khoekhoe *†namib* ~ *†namis* 'simple hand-made knife' is related to *†nami*, a root occurring in the Damara dialect and meaning 'to chip, chop off' inter alia. Koyraboro Senni *huri* also dialectally means 'seek, look for,' while *zaama* dialectally is also a particle meaning 'because.' Ngambay *ḳiya* also means 'to hide,' 'to place, set.' The Burarra term *angujarrcha* contains *jarrcha* 'to slice off,' while Dadibi *ge hwa* contains *ge*, meaning 'nut, egg' or 'small object' in general. Ngaanyatjarra *kunmarnu* is an avoidance register term for 'Sunday, week, jumper,' 'key,' and 'knife.' Rotokas *visi-paa* is analyzable as /visi-pa/ 'poke/hit-DERIV.' Toaripi colexifies "tying, lashing material" inter alia, and Badaga *su:ri* also means 'sharp, pointed' in an adjectival sense. Japanese *hō-chō* is analyzable as 'kitchen-man.' Haida *q'it'uhl 'laʔáaw* contains the instrumental prefix *q'it'* 'cut with knife' and presumably *uhl* 'shape.' Itzaj *lomik* contains *lom* 'to stab,' and also denotes a 'stab,' 'stake,' or 'nail.' The Kashaya term *qahca* also means 'missile' and 'clitoris.' Nuuchahnulth *ʕak-ýak* is analyzable as /ʕakʷ-ýak/ 'whittle-tool,' and Wintu *čebet* 'stone knife' is related to *čEb-* 'sharp, knifelike, plane, whittle.' The Arabela term *cushiishi* contains *cushi* 'pig' (perhaps folk etymology of Span. *cuchillo*?). Maxakalí *mikax* also means 'rock, stone,' Bislama *naef* also 'blade' specifically, and Rotuman colexifies 'knife' with 'to circumcise' inter alia. Samoan has *fa'aola fanua* 'save/savior land/field' for 'adze, axe, knife' (the term is restricted to polite usage), and Sedang *řokong* is also used with the meaning 'mouth, language, word.'

84. *The Ladder*

Representation: 78%

Motivated: 52.9%

Thereof Analyzable: 33.3%

Thereof Colexifying: 20.0%

Thereof by Contiguity: 24.2%

Thereof by Similarity: 25.4%

Recurrent associated meanings: stairs/staircase, climb/ascend, step, bridge, wood/tree, foot, scale, road/path, leg, hole, thing, sky, walk, lean

Frequently, terms for this meaning are of the derived kind, the derivation base being verbs meaning ‘to climb,’ ‘go up,’ or ‘ascend,’ as in Rotokas *iipa-pa* ‘climb/go.upward-DERIV.’ Alongside semianalyzable terms in Ineseño Chumash, Hani, and Kapingamarangi, this is realized by derived terms also in Yoruba, Muna, Sora, Blackfoot, Upper Chehalis, Carrier, Cheyenne, Chickasaw, Kiliwa, Nez Perce, Wintu, Central Yup’ik (colexifying “neck opening of parka”), Aguaruna, Bora, Chayahuita, Guaraní, Fijian, and Rotuman. There are also terms with an additional element bearing lexical meaning. Mbum has *fè-hénà* ‘thing-climb’ (a similar term with ‘thing’ is also featured in Hupda), Khoekhoe has *!apa-haib* ‘climb/ascend-stick/tree’ for a traditional type of ladder. Somewhat differently, Carrier has *ukwe-tera-î-thi* ‘on.it-getting-up-road’ (complex terms with ‘path’ are also Lengua *amai letin* /amai netin/ ‘path sky/above,’ which colexifies ‘ladder’ with ‘dam’ as well as Hawaiian *ala-pi’i* ‘path-climb,’ which denotes ‘ladder’ as well as ‘step,’ while the association is realized by colexification, also with ‘door’ and ‘gate,’ in Haida). Moreover, Piro has *hatsko-pi-xe* ‘ascend-rod-pole,’ and similarly, Blackfoot has *iihtáísokamisáóóp* /iiht-á-sok-wamis-oo-o’p/ ‘INSTR-DUR-above-??-go-21.NOM.’ This term, like the Hawaiian one, colexifies ‘ladder’ with ‘step,’ and this is indeed a very frequent association in the languages of the sample, and some of the terms derived from ‘to climb’ above exhibit this pattern as well. It is also found in Dongolese Nubian, Kaluli, Lavukaleve, One (here also with the meaning “horizontal brace of a pangal bed”), Sahu, Toaripi, Khalkha (also with ‘footboard’ and ‘pedal’), Upper Chehalis, Arabela, Cayapa, Bislama, Fijian, and Sedang, which also colexifies ‘stem’ (the original meaning of Japanese *hashigo* is also ‘step’ diachronically). Similarly, Kaingang has a term for ‘ladder’ derived from a verb meaning ‘to step.’ A semianalyzable term for ‘ladder’ where one of the constituents means ‘step’ is found in Wayampi, and one where it means ‘to step on, set foot on’ in Kwoma. Another common pattern of colexification is that with ‘stairs’ and/or ‘staircase,’ found in Efik, Dongolese Nubian, Yoruba, Berik, Muna, Rotokas, Sentani, Toaripi, Basque, Khalkha, Blackfoot, Upper Chehalis, Chickasaw (which also colexifies ‘fire escape’ specifically), Comanche, Nez Perce, Pawnee, Central Yup’ik (colexifying also ‘rung’), Embera, Jarawara, Maxakalí, Hani, Hawaiian, Kapingamarangi, Manange, Rotuman, Bislama, Fijian, Manange, and Mandarin.

Furthermore, Basque has *esku-eskailera* ‘hand-stairs,’ and Khalkha *giški-gyr* ‘step-bridge,’ ‘bridge’ (or specific types of bridges) and ‘ladder’ are furthermore colexified in Carib (which also colexifies ‘harbor’), Lengua, Toba, Yanomámi (where the relevant term *ihiraki* is derived from *ihira* ‘to construct a frame’ by means of suffixation of the quantal classifier *-ki*, for which see § 4.4.1.1.), Hawaiian (‘plank bridge’ more specifically, and also colexifying ‘trestle’), and Sedang. Five sampled languages have complex terms in which

one of the constituents means 'foot': Efik *udik'uküt* /u-dik'hi-uküt'/ 'NMLZ-tread.upon-foot/leg,' Abzakh Adyghe *leywen* /λ(e)-ye-we-n/ 'foot-UPWARD.MOTION-stomp-HAVE.INTENTION.TO,' Kiowa *'ān-t'out-'H'dā* containing *'ān* 'foot' and *'H'dā* 'pole' (compare also *'ān-t'out* 'to climb up steps'), Miskito *minamangka*, containing *mina* 'foot' and *mang* 'to put' (there is a further semianalyzable term with *mang* in this language), and Bwe Karen *kha-bó* 'foot/leg-handle/holder.' Further complex terms involving 'wood' or 'tree' are Biloxi *a'ya i'de' ~ aya'i'de /a'ya* i'de/ 'wood INSTR-go,' Tuscarora *urē?náhrareh*, consisting alongside grammatical material of the verb *ahrār*- 'be a hole' and the incorporate *-rēT*- 'tree, log' (compare colexification of 'hole' and 'ladder' inter alia in Hawaiian), Maxakalí *mīp-ku'in* 'wood-slashes/stripes,' and Ket *bulaŋdoks* /būl-aŋ-d-ōks/ 'leg-PL-POSS-wood' (for the association with 'leg' compare also Efik *uđiri-uküt* 'top-leg'). Note also that the Cashahuita term is derived from a verb meaning 'to climb' by a classifier for wood-related items, and that there are semianalyzable term with 'wood' in Basque and Aguaruna.

Efik *ē-beri* is derived from *beri* 'to lean,' and a term where the meaning 'leaning' figures is also found in Pawnee (an optional complex term with a verb meaning 'to lean up' is also found in Haida on the basis of the term colexifying 'door,' 'gate,' and 'ladder' mentioned above). Alongside Lengua *amai letin* 'path sky/above,' which was already mentioned above, Kiliwa also has a term betraying an association with sky: *?-mai?=t-h-?+paa-y-u?* is analyzable as 'DN-sky/heaven=SUBJ-3-DN+depart-ATT-OBL.' Dadibi *togobili* may be analyzable as /togobe-bilibo/ 'edgeposts-walk,' and Abipón *n-acaR-haR-late* as 'POSS.INDEF/3SG-walk-the.one.who-LOC.' This term also means 'shoe' (compare section 91). Finally, Itzaj, Embera, Fijian, and Hawaiian colexify 'ladder' with 'scale,' and Rotuman colexifies 'ladder' also with 'stave, staff' and also music written in this notation, and Hawaiian with 'scale' in music.

Other associations include: Buli *tiili* colexifies 'ladder' with "the clay connection between two ancestral shrines." Hausa *tsani* also denotes an 'intermediary' inter alia, and Kanuri *kùrángá* also means 'monkey.' Ngambay *dáiki* is also used to refer to "something difficult," and *mbata* colexifies 'stool.' Yoruba *àkàbà* contains *bà* 'perch on, alight' and the nominalizer *á-*. Baruya *kwaaka* also means 'soil, earth, dirt, clay, ground, land, country' as well as "shooting short of a target," while Gurindji *tankuj* also may refer to a 'useful thing.' Kwoma *akatoko* (containing *aka* 'house') is also used with the meaning 'escalator,' and *piitiishey* also means 'scaffolding.' Muna *lawa* also means to "answer back, respond" as a verb, and for *pulangku ~ polangku*, compare *langku*, which denotes a part of a loom, but also means "social rank, level." Ngaanyatjarra *lata* also means 'letter' (due to collapse of Engl. *ladder* and *letter*). Kosarek Yale *modobak* is derived from the verb *modob*- 'walk or climb using footholds' and colexifies 'foothold' as well as 'pole with notches.' Badaga *so:pa:na* is also the name of a "log in which steps have been cut," Khalkha *šatu(n)* also means "phase, stratum, level," and Welsh *ysgol* also means 'school.' Oneida *yelathastákhwa?* is analyzable as /ye-lathast-hkw-ha?/ 'FEM.INDEF.SG.AGENT-get.something.up-INSTR-HAB,' Central Yup'ik *akeq* colexifies 'ladder' with 'barb' and 'rung,' and Bora features also a term on the basis of a verb meaning 'to descend' rather than 'to ascend' as reported above: *niityé-wááhyo* is analyzable as 'descend-CL.layered.things.' Cashinahua *tapaiti* is derived from *tapa* 'floor' by means of the instrumental suffix *-ti*. Jarawara *ki-kisima* is analyzable as 'RED-come.down.'

Toba *piaxalate* ~ *napiaxalate* also means ‘pulpit,’ and the Tsafiki term *teranca* appears to contain *terano* ‘to dance’ and *ca* ‘in front of.’ Bislama *step* also conveys the meanings ‘terrace on cliff’ and ‘gait, pace,’ while Hawaiian *haka* also means ‘platform, shelf’ *inter alia*, and Manange *2li* also ‘face.’

#### 85. *The Mirror*

Representation: 77%

Motivated: 51.3%

Thereof Analyzable: 34.5%

Thereof Colexifying: 16.8%

Thereof by Contiguity: 24.5%

Thereof by Similarity: 5.3%

Recurrent associated meanings: glass/type of glass, see, look/inspect, shadow, window, reflect/reflection, water, face, eyeglasses, eye, bright/brighten, thing

As for virtually all artifacts, terms making reference to the purpose they serve abound. In this case, this means that very frequently, terms make reference to seeing, looking (see Buck 1949: 454 for Indo-European), or reflecting. Derived terms on the basis of verbs meaning ‘to see’ frequently also contain a reflexive marker (as did a Sanskrit term, Buck 1949: 454), as in Chickasaw *aailipisa*, analyzable as /aa-ili-pisa-’/ ‘LOC-REFLX-see-NMLZ.’ There are also derived terms in Mali (where the meaning of the derivation base is ‘see past’ more precisely), Upper Chehalis, Comanche, Kashaya, and Nez Perce. In lexical terms, ‘shadow’ is often the meaning of the additional lexical element (an association likewise reported for Indo-European by Buck 1949: 454), as in Mbum *ákó-têm* ‘see-shadow,’ and also in Japanese and Pawnee (associations with ‘shadow’ also occur in other configurations, by direct colexification in Lavukaleve, here also with ‘spirit,’ Rama, also with ‘picture,’ as well as Sedang, and by the analyzable Chukchi term *wiilyiteney* containing *wiil* ‘shadow’ and *yite* ‘watch’ and Samoan *fa’a-ata* ‘CAUS-shadow/image’). Otherwise, Efik has *u-kur-isü* ‘NMLZ-see-face,’ and Yir Yoront *kowllewkarrlnh* and *kowllewkerr(w)lh* contain *kowllew* ‘face’ and *karr* ‘see, look at’ and furthermore colexify ‘mirror’ with ‘photograph.’ Biloxi *o”do”ho”ni* revolves around the root *do* ‘to look at, see,’ Cheyenne *amôhóomâhtsestôtse* contains *amôhóom* ‘to see in reflection’ and *hestôtse* ‘thing’ (and indeed also means ‘reflection’), and semianalyzable terms with ‘see’ occur in Buli, Hausa, Tehuelche, and Great Andamanese. Terms derived from verbs meaning ‘to look’ or ‘inspect’ are found in Kanuri, Rotokas, Kiowa, Oneida and Fijian, where the derivation base means ‘to look at, as a reflection in water or in a mirror’ specifically. Rotuman colexifies ‘mirror’ with “to watch closely, gaze at” directly. Furthermore, Yoruba has *à-wò-jijì* ‘NMLZ-look-shadow/reflection,’ Nivkh *un’yr-njus* ‘star-place.to.look,’ and Carrier *pê-na-tse-n-de-nel’ên* is analyzable as ‘wherewith-repeatedly-?-rotundity-at.one’s.own-look,’ with the element glossed as ‘rotundity’ referring to the ‘face’ (other complex terms more or less tightly connected with ‘face’ other than those already mentioned are Itzaj *eetz’-ich* ‘face-make’ (‘face’ in the sense of ‘grimace,’ though), Cashinahua *beisikiti*, presumably analyzable as /beisikiki-ti/ ‘look.at.other.person’s.face-INSTR’ and Cubeo *jiva-rû* ‘face-CLASS.ROUNDISH.OBJECT’). Terms derived from verbs meaning ‘to reflect’ or nouns meaning ‘reflection’ are found in the Nunivak Island dialect of Central Yup’ik, where the relevant term *tarenriurun* is analyzable

as /tarenriur-(u)n/ 'look.at.one's.reflection-device.for,' Malagasy (*fi-tàratra* 'NMLZ-light.beam/reflection'), and also in Haida and Tetun, while Kashaya *ʔama puṭʰam* contains *ʔama* 'thing' and *ṭʰaṵm* 'to reflect;' Wintu *ʔilʔiloqma* can also refer to "anything that reflects; something shiny."

A pattern that is also widespread in the languages of the world is colexification of 'mirror' with 'glass' or types of glass, such as colexification of 'mirror' with 'pane of glass' specifically in Swahili. Such associations occur also in Hausa, Yoruba, Burarra, Kyaka, Muna, Ngaanyatjarra, Abzakh Adyghe, Bezhta, Upper Chehalis, Highland Chontal, Itzaj, Pawnee, Tuscarora, Great Andamanese, Fijian, Hawaiian (also colexifying 'clear, transparent, obvious' and 'cool' among other meanings), Bwe Karen, Malagasy, Sedang, White Hmong, Rotuman, Takia, Yay, and Bislama (note that according to Buck 1949: 454, this pattern within Indo-European "seems to be peculiar to English"). White Hmong, in addition, has *tsom-iav* 'look-glass.' Due to this general pattern of colexification, Hausa, Bislama, and White Hmong also colexify 'mirror' with 'eyeglasses.' Furthermore, Upper Chehalis, Kashaya, Kiowa, Pawnee, Tuscarora, Wintu (by the term *kenwi-na's*, containing *ken* meaning inter alia 'in, be in, put in' and presumably *wi* 'male;' there is a reconstructed root \**wi* meaning 'person' more generally), and Fijian colexify 'mirror' with 'window,' and there are other similar scattered patterns of colexification, such as for instance that with 'telescope,' 'binoculars,' 'thermometer,' and 'barometer' in Bislama, which also colexifies 'car window' specifically.

Another rather unrelated pattern is the association between 'mirror' and 'water,' due to the reflecting properties of water. This association, in the form of a pathway of semantic extension under circumstances of acculturation, is hinted to by the fact that Tuscarora *awé'kyeh*, meaning 'glass,' 'window,' and 'mirror,' also means 'liquid.' Fijian also directly colexifies 'mirror' with 'water' by the analyzable term *i iloilo* 'DERIV look.at.reflection-RED,' Santiago Mexquititlan Otomí colexifies 'mirror' with 'river,' Noni has *me ε joo* '1SG in water' (the term also denotes a 'riddle'), Angkor *hoe-himbo* 'water-ear,' Toaripi *ma ove* 'water picture' (lexicographer's comment: "before glass or metal mirrors were available, reflections in water served as a mirror"), and Kiliwa *xaʔ=ny-ʔ+yuw-l=t-p-sa-w-uʔ* 'water=POSS-DN+eye-ILL=OBJ-MP+see-DIR-OBJ.' In Buin, *tiinura* is a term for "water collected in trees" that is used as a mirror, and Berik has a semianalyzable term. The association with 'eye' as found in Kiliwa is also present in Ket (*deskənejroks* /dēs-kənej-r-oks/ 'eye-ʔʔ-POSS-wood'), and semianalyzable terms where one of the constituents is 'eye' are featured in Waris (where this term is archaic) and Badaga. Finally, Kashaya has *ʔama piṭʰam*, containing *ʔama* 'thing' and *-ṭʰaṵl-* 'bright, light enough to see by,' and similarly, Yanomámi has *warara-rimi* 'brighten.up-NMLZ'

Other associations include: Hausa *madubi* and *soka* also mean 'sorcerer' and 'wizard,' Muna *paeasa* assumes the meaning to "follow the example" when used verbally, and Basque *ispilu* also means 'white spot, white patch.' Khalkha *toli(n)* is also used with the meanings 'metal plate' and 'dictionary, vocabulary,' and Welsh *drych* also means 'sight.' Blackfoot *sáapia'tsis* contains *sap* 'in, within' and the instrument nominalizer *-a'tsis*, Kiliwa *xʔ=al=h-haʔ-uʔ* is analyzable as 'CAUS+ʔʔ=3-see-OBL,' Nuuchahnulth *piihyaksata* as /piḥ-ýak-sata/ 'observe/study/examine-tool-at.or.on.the.forehead,' while Santiago Mexquititlan

Otomí *hñe* also means ‘to put on’ and San Lucas Quiaviní Zapotec *gyigwàan* also ‘very beautiful girl.’ Piro *haniha-pi* is analyzable as ‘shade.eyes-rod,’ and Wichí *to-peyak-hi* as ‘POSS.INDEF-image-LOC.in,’ while Hawaiian *aniani* appears to be reduplicated from *ani*, one of the meanings of which is ‘for a hand to pass over a surface.’

#### 86. *The Needle*

Representation: 90%

Motivated: 35.8%

Thereof Analyzable: 14.5%

Thereof Colexifying: 21.6%

Thereof by Contiguity: 12.2%

Thereof by Similarity: 11.0%

Recurrent associated meanings: injection/syringe, sew, pin, thorn, nail, hand of clock, thread/string, bone, cloth, fork, net shuttle, pierce, thing

For ‘needle’ (or ‘awl’), the most frequent association is unsurprisingly that with ‘to sew’ (see also Buck 1949: 412), by terms of the derived type for in Central Yup’ik (which has, for instance, *mingqun* /mingqe-(u)n/ ‘sew-device.for’), Upper Chehalis, Chickasaw, Kiliwa, Arabela, Fijian (where *cula* not only means ‘to sew’ but also ‘to pierce, let blood, vaccinate,’ for this compare Nuuchahnulth *qačak* /qač-ak<sup>w</sup>/ ‘pierce/puncture-tool’), and Malagasy. Terms of the lexical kind have ‘thing’ as the meaning of the additional constituent in Koyraboro Senni *taa-haa* ‘sew-thing’ (which is a substitute term for a monomorphemic term for ‘needle’ that must not be uttered at night), ‘cloth’ in Comanche *wana tsahkuna?* /wana tsahkunarə-?/ ‘cloth sew-NMLZ,’ and ‘bone’ in Kapingamarangi *iwi due mee* ‘bone sew thing’ (semianalyzable terms containing verbs meaning ‘to sew’ are attested in Nuuchahnulth, Miskito, and Tehuelche, and there is one containing ‘thing’ in Dadibi). Both associations just mentioned are also found in different configurations. Buli has *garupein* /garuk-pein/ ‘cloth-arrow,’ Yir Yoront *mipkallnh*, containing *mip* ‘cloth’ and *kal*, meaning inter alia ‘spear, poke, pierce, cut,’ Rotuman *sui susuag ha’u* /sui susuaga ha’u/ ‘bone sewing clothes,’ and a semianalyzable term where the identifiable constituent can be identified to mean ‘veil’ is found in Greek. Wappo and Yuki exhibit the association with ‘bone’ by colexification. In Upper Chehalis and Rama, ‘needle’ and ‘thread’ are colexified (in Rama ‘silk grass needle’ more specifically), while Lengua has *sokyi-tama* ‘carry-string’ and Piro *tsa-pu* ‘thread-ENTITY.’ Similarly, Hawaiian colexifies ‘needle’ with verbal ‘to thread beads’ and ‘to string pierced objects.’ Presumably by provenience contiguity, Anggor, Baruya, Buin, Khalkha, Abipón, Bora, Chayahuita, Tehuelche, and Yanomámi colexify ‘needle’ with ‘thorn’ (Yanomámi also with ‘porcupine quill’), and Hupda has *măc-?ut* ‘metal-thorn.’ Samoan colexifies ‘needle’ with ‘sting’ (see Buck 1949: 412 for cognates evidencing this association by semantic shift in Indo-European). There are also patterns of colexification due to semantic extension. Basque, Khalkha, Lesser Antillean Creole French, and Tetun colexify ‘needle’ with ‘hand of clock,’ and Koyraboro Senni, Ngambay, Noni, Rendille, Swahili, Anggor, Baruya, Bezhta, Khalkha, and Cubeo with ‘injection, syringe.’ Kwoma, Rotokas, Oneida, Tuscarora, Miskito, Tehuelche, and Hawaiian colexify ‘needle’ with ‘nail’ (Oneida also with ‘wire,’ and Tuscarora, Miskito, and Hawaiian also with many other objects made of metal, in Tuscarora for example ‘auger,’ ‘fork’ – this pattern of colexification is shared

with Nez Perce and Nuuchahnulth – ‘key,’ ‘pitchfork’ and “dwarf pike”), and Xicotepec de Juárez Totonac *līxtocon* is analyzable as /lī-xtokó-n/ ‘INSTR-to.nail-INSTR.’ Koyraboro Senni, Rotokas, Basque, Ket, Kiowa, Oneida, Cayapa, Hawaiian, and White Hmong colexify ‘pin,’ and San Mateo del Mar Huave and Nez Perce colexify ‘needle’ with ‘net shuttle.’

Other associations include: Ngambay *énemé* also denotes a “kind of drill to make a hole with,” and *suwa* also means “boil leaves rapidly in water, cook meat.” Kyaka *wamyali* ~ *wamalyi* also means ‘skewer.’ Basque *orratz* is also used with the meanings ‘needle of grammophone,’ ‘compass,’ and ‘switch,’ while the denotational range of the Ket term in also includes ‘fingernail’ and ‘claw.’ Khalkha *zegyn* ~ *zegyy* also means ‘East, oriental, left,’ and Japanese *hari* also means ‘beam,’ with the terms being prosodically different. Biloxi *α"sadûki* contains the word for ‘pine tree,’ whereas Cheyenne *vé'ho'êškóovo'hestôtse* is analyzable as /vé'ho'ê-škóovo'hestôtse/ ‘whiteman-toothpick.’ San Mateo del Mar Huave colexifies ‘needle’ with ‘snake, worm,’ Itzaj *puutz'* also means ‘enemy,’ and Kiliwa *t-x-pa?-p-u?* is analyzable as ‘OBJ-CAUS-possess.round.object-MP-OBL.’ Wintu *čurp* also denotes the concept ‘dagger’ and anything with a sharp point in general. The Yaqui term *ji'ikiam* is analyzable as /ji'ik-ia-im/ ‘weave-NMLZ-PL.’ Abipón *n-icir-en-kate* (variant form *nicirenRat*) is analyzable as ‘POSS.INDEF/3SG-unite-VOL-INSTR,’ and Bororo *akigu ičira* as ‘thread palm.fibre.’ Hawaiian *pahēle*, also denoting a ‘snare, noose’ or ‘trap’ as well as ‘deceit, treachery,’ seems to be derived from *hele*, meaning ‘to tie, bind, lash, snare, noose’ inter alia. Another Hawaiian varies somewhat in form depending on the variety of Hawaiian; that spoken on the island of Hawai'i, *hānai*, also means ‘foster child, stepchild’ inter alia. Finally, Manange *4<sup>th</sup>e* also means ‘to hear,’ Mandarin *zhen*<sup>1</sup>, going back to distinct Early Middle Chinese terms (Pulleyblank 1991: 401), also “true, exactly,” and Yay colexifies ‘needle’ with ‘gold.’

### 87. *The Paper*

Representation: 84%

Motivated: 59.6%

Thereof Analyzable: 13.2%

Thereof Colexifying: 46.8%

Thereof by Contiguity: 48.0%

Thereof by Similarity: 5.2%

Recurrent associated meanings: letter/book/document, write/draw, leaf, talk/speak, playing card, cloth, map, mail, role, cardboard, banknote, skin, wrap/wrapping

The most frequent association is, by contiguity, colexification with ‘letter,’ ‘book,’ or another type of document and sometimes also ‘page.’ It is found in Buli, Efik, Hausa, Koyraboro Senni (colexifying also ‘amulet, talisman’), Mbum, Noni, Dongolese Nubian, Rendille, Yoruba, Buin, Burarra, Kwoma, Yir Yoront, Abzakh Adyghe, Badaga, Basque, Ket, Biloxi, Cahuilla, Upper Chehalis, Cheyenne, Chickasaw, Highland Chontal, Comanche, Haida, San Mateo del Mar Huave, Itzaj, Kashaya, Lake Miwok, Lesser Antillean Creole French, Nez Perce, Oneida, Santiago Mexquititlan Otomí, Pawnee, Pipil, Tuscarora, Central Yup'ik, San Lucas Quiavini Zapotec, Abipón, Aguaruna, Bora, Bororo, Cashinahua, Cavineña, Cayapa, Chayahuita, Guaraní, Huambisa, Hupda, Kaingang (where *vēnh ra* seems to contain *vēnh* ‘small plants’), Macaguán, Maxakalí, Miskito, Piro, Rama, Sáliba, Wayampi,

Yanomámi, Bislama, Hani, Malagasy, White Hmong, and Samoan. The relationship between ‘letter’ and ‘paper’ is marked by gender alternation in Embera, and that between ‘sheet, piece of paper’ and ‘book, letter’ in Khoekhoe by alternation of nominal designants. More specifically, Greek, Hawaiian, and Samoan colexify ‘paper’ with ‘playing card’ (Samoan also with ‘banknote’ and ‘ticket’), Greek and Haida with ‘map,’ Nez Perce and Central Yup’ik with ‘mail’ (the Nez Perce term also denotes the ‘Bible’ specifically), Basque and Itzaj with ‘role,’ Abzakh Adyghe and Samoan with ‘banknote,’ and Cashinahua and Rotuman with ‘cardboard.’

Otherwise, complex terms of the derived type from a verb meaning ‘to write’ or ‘to draw’ occur in Efik (*ñ-wet* ‘NMLZ-write/mark/paint;’ this term also denotes a ‘printing, inscription’ and ‘form,’ ‘impression,’ ‘representation,’ ‘reflection,’ and ‘shadow’), Sora (*id’o:l-ən* ‘write-N.SFX’), Arabela (*naajio-jua* ‘writing-CLASS.GROUND;’ this term also denotes a ‘scribe’ and a table used to write on), Abipón (*elerk-a* ‘writing/letter-PL’), and Yanomámi. Indeed, the Abzakh Adyghe and Oneida terms colexifying ‘paper’ with ‘letter’ and ‘book’ are also of this type (the Chukchi, Cheyenne and Yanomámi colexifying terms are semianalyzable). The Yanomámi term colexifies ‘paper’ with ‘pen.’ Moreover, Tsafiki has *pilá quidó* ‘writing/book skin/bark’ (note also the colexification of ‘skin,’ ‘hide,’ and ‘paper’ in Buli and the origins of Indo-European words for ‘paper,’ Buck 1949: 1289), and Chickasaw *holisso* also means ‘to be written’ in verbal usage. Kwoma, Rotokas, Bezhta, Hupda, and Piro colexify ‘paper’ with ‘leaf’ (Kwoma with “dry banana leaves” specifically), as did Sanskrit (Buck 1949: 1289); note also Kaluli *mo:fo:s*, containing *fo:s* ‘leaf’ and perhaps *mo:* “base of tree stump or trunk,” “basis or reason for utterance.” In two languages of Eastern North America, there are complex terms for ‘paper’ where one of the constituents means ‘cloth:’ Lakota *mniñúha-khakháka* ‘cloth-rustle’ and Pawnee *raawihaakaraa’iit*, which is analyzable as /*raawir-taakaar-raa’iit*/ ‘cloth-white-telling.’ Moreover, Nuuchahnulth colexifies ‘cloth’ with ‘paper;’ the relevant term is *qicaaʔ*, analyzable as /*qic-ǎʔ*/ ‘mark/paint/tattoo-on.a.fabriclike.surface.’ The Pawnee term, betraying an association with uttering words, has a parallel in Nez Perce (*tí’mie-s* ‘speak-AGT’) and in Abipón, where there is a word for ‘paper’ (as well as ‘word’) derived from a verb meaning ‘to talk’ (see also Buck 1949: 1003 for this association in Ancient Greek). Baruya colexifies ‘paper’ with ‘wrapping,’ and, analogously, Kiowa *mą́tṣḥę-mą́* ‘sheet of paper’ is analyzable as ‘wrap-NOUN.POSTFIX.’ Bororo *bqpera* also denotes leaves from the stomach of ruminants; Kiliwa *txpha?* may also perhaps refer to a “part of the ruminant digestive system called ‘the book’ (Sp. el libro), or related to sheepskin parchment,” and similarly, Lake Miwok *pápel* also denotes “an internal organ attached to the stomach of animals which opens like a book when one cleans it.” Since the Lake Miwok term is a borrowing from Spanish and the Kiliwa gloss refers to Spanish, it seems likely that this is a pattern copied from (local) Spanish.

Other associations include: Bakueri *liwawé* also means ‘wing,’ Buli *gbang* also ‘gambling,’ and Rendille *khadaab* also ‘scab.’ Burarra (-)jurra colexifies ‘paper’ with ‘track, footprints,’ and Kyaka *pepa* may also refer to ‘stationry’ or ‘tissue.’ Yir Yoront *waqrr* also denotes the bark of the tea tree or melaleuca, and Japanese *kami* also denotes the ‘hair’ and ‘God,’ with the individual meanings distinguished in writing. Biloxi *akūtxyi’* contains the word for ‘spotted, striped,’ and the Carrier term *etestles* also means ‘fur.’ Pipil *a:mat* is



also the name of the amate, a species of fig tree, Central Yup'ik *igaq* also means 'mark, symbol' and 'letter of alphabet,' and Bora *waajácuháámi* is analyzable as /waajácu-ʔa:mi/ 'study/know/understand-scm.leaf.' Rotuman *pepa* also means 'pepper' (< Engl. *paper* and *pepper* respectively).

#### 88. *The Pen*

Representation: 66%

Motivated: 51.6%

Therof Analyzable: 45.4%

Thereof Colexifying: 6.2%

Thereof by Contiguity: 39.7%

Thereof by Similarity: 3.4%

Recurrent associated meanings: write, stick, feather, paper, wood/tree, crayon, ink, rod, chalk, paint, brush, typewriter, thing, poke, bone

Terms for 'pen' frequently contain verbs meaning 'to write,' which may in turn themselves have experienced semantic extension from 'to scratch,' 'to mark,' or like meanings which are not discussed separately here (the association between 'pen' and 'writing' is quite rare in the diachrony of Indo-European, mentioned only for Sanskrit by Buck 1949: 1290). Derived terms, such as Blackfoot *ihtáísínaakio'p* /iɪht-á-sínaaki:ʔʔ-o'p/ 'INSTR-DUR-write:ʔʔ-21.NOM' are featured in Rotokas, Upper Chehalis, Chickasaw, Kashaya, Lake Miwok, Nez Perce (colexifying 'seal'), Oneida, Quileute, Xicotepec de Juárez Totonac, Tuscaraora, Yuki, Central Yup'ik, Aguaruna, Arabela, Bora, Carib, Cashinahua, Cubeo, Guaraní, Kaingang, Piro, Imbabura Quechua, Toba, and Yanomámi. Variants of this pattern are found in Haida, where the derivation base of a relevant term (colexifying 'chalk' and 'crayon,' a pattern of colexification shared with Piro, and with Lesser Antillean Creole French in the case of 'crayon') means 'to write a letter' specifically. Cheyenne and Mbum have analyzable terms of the lexical type where the second constituent means 'thing.' Otherwise, 'wood' or 'tree' is a frequent meaning for the second constituent, as in Pawnee *raktariihkaruukus*, analyzable as /rak-rariihkara'uk-hus/ 'tree/wood-write-IPFV.' Such terms are also found in Ngambay (where 'tree, wood' is colexified with 'stick' inter alia), Kiowa (colexifying 'writing table'), Maxakalí, Tsafiki (in the latter two languages, 'tree' is colexified with 'bone'), and Yir Yoront, where the relevant term *yoqlatkallnh* contains *yoq* 'tree, stick, wood,' *lat* 'paper' and *kal*, meaning 'spear, pierce, poke, cut' inter alia (the association with 'poke' may be an Australian pattern, compare Gurindji *turrrp-turrrp-kaji* 'poke-RED-AGENT'). In addition, Comanche has *parúbooʔ* /paa-túboorɘ-ʔ/ 'water-write-NMLZ' for an 'ink pen' specifically, and Abzakh Adyghe *šherə-tʃe* 'with.head/upperside-write.' In Hani, there is a semianalyzable term also denoting a 'writing brush' (a meaning also associated with 'pen' in Mandarin) with the identifiable constituent meaning 'to write' and 'to rot, decay' (the other one, *dul*, means 'similar, true to life' when occurring on its own.)

Sora *ido!kappa:n* contains '*kappa:-* wing' alongside *id'o:l-* write,' and Khoekhoe *xoa-lam-mi* is analyzable as 'scrape/write-feather-3SG.MASC.' Both point to a pattern particularly common in the Old World (as also evidenced by its frequency in Indo-European, Buck 1949: 1290): since quills were once commonly used as a writing instrument, Basque, Nivkh, and Kildin Saami colexify 'pen' with 'feather' and/or 'quill' (the Basque term *hegats*

also means ‘fin’ and ‘eaves’ and may be related diachronically to *hatz* ‘finger,’ by another term Basque also colexifies ‘snowflake’), while Efik has *ntañwet* /ntañ'-ñwet/ ‘feather-paper’ and Hawaiian *hulu kākau* ‘feather write.’

In turn, ‘paper’ (which potentially colexifies certain types of documents inter alia, cf. section 87) is of course by functional contiguity frequently associated with ‘pen,’ as already seen in the Yir Yoront and Efik terms mentioned above. ‘Paper’ and ‘pen’ are colexified in Yanomámi (similarly, Huambisa colexifies ‘pen’ with ‘booklet’). Furthermore, Bororo has *bapera atugo epa* ‘paper writing instrument,’ Biloxi *akütxyi' o"ni ~ akütxyi' o"* ‘paper make,’ Carrier *testles-tcen* ‘paper-stick,’ and Miskito *ulb-aia dusa* ‘write-INF stick’ (compare also Baruya *pikariyita* /pikarya-yita/ ‘carving-stick,’ colexification of ‘log, rafter 2x4, a limb, a young tree, a smaller log, any piece of wood that is like a stick’ with ‘pencil’ in Wintu, and the fact that the Bora term is derived from a verb meaning to ‘write’ by a classifier for small sticks). Rather than ‘stick,’ the second element is ‘rod’ in Piro and Imbabura Quechua; due to colexification with ‘tree, wood’ and ‘stick,’ this association is also present in Ngambay. Another complex term of the lexical type where one of the constituents is ‘write’ is Santiago Mexquititlan Otomí *xini t'ot'i*, containing *xi* ‘hair’ and *ot'i* ‘to write.’ White Hmong has *cwj-mem* ‘stick-ink,’ a derived term from a base meaning ‘ink’ is also found in Central Yup'ik (*ingeg-cuun* ‘ink-device.for’), and a term where one of the constituents is ‘ink’ is featured in Yay. Rotokas and Aguaruna colexify ‘pen’ with ‘typewriter’ (Aguaruna also with ‘booklet,’ ‘notebook’), and Bislama and Takia with ‘paint,’ in both cases due to collapse of English source words *pen* and *paint*.

Other associations include: Hausa *alkalami* also can refer to ‘a digit in arithmetic’ inter alia, Muna *koroka* denotes a fibre rib of the sugar palm which is used as a pencil, Khalkha has *zyyg*, which is also a variant of *zyyg ~ ysyg* ‘writing, letter.’ Welsh *pin* also means ‘bobbin,’ and Japanese *fude* is analyzable as /fumi-te/ ‘text-hand.’ Nivkh colexifies ‘pen’ and ‘pointed drill.’ Hawaiian *peni* (< Engl. *pen*) has fallen together with English *penny*, which was also borrowed, and Rotuman *pene* indigenously also means “to emit an odour.”

#### 89. *The Rope*

Representation: 86%

Motivated: 55.5%

Thereof Analyzable: 5.1%

Thereof Colexifying: 50.4%

Thereof by Contiguity: 5.1%

Thereof by Similarity: 32.4%

Recurrent associated meanings: thread/string/cord/twine, vine/climbing plant, tie, line, lasso, strap, fishing line, badge of rank, cable, belt, bundle of rice, thong, whip, umbilical cord

The typical structure for many artifact terms - derived terms from the associated action that can be performed with the artefact - is relatively rarely found for this meaning. Muna *ka-tapu* is analyzable as ‘INSTR-tie/tether,’ and such derived terms are otherwise only found in Khalkha (where the derivative also means ‘hitching post’ and ‘training a horse for a race’), Chickasaw, Central Yup'ik, and Tehuelche (see Buck 1949: 550 for evidence from Lithuanian as well as from cognates meaning ‘cord, band’ and ‘bind’ respectively in San-

skrit and Greek). Kyaka, Khalkha, and Wintu terms directly colexify 'rope' and 'to tie' (and there are sometimes redundant complex terms on the basis of the relevant root). More frequently, namely in 20 languages (Ngambay, Baruya, Buin, Kwoma, Kyaka, Mali, Rotokas, Sahu, Kosarek Yale, Sora, Wintu, Bora, Chayahuita, Hupda, Tsafiki, Wayampi, Fijian, Sedang, Takia, and Bislama), is colexification with (a specific kind of) 'vine' or 'climbing plant' generally, either by perceptual similarity or, more likely, by provenience contiguity (see Laycock 1970: 1160 for New Guinea specifically). Very frequent is colexification with 'thread,' 'string,' 'cord,' and/or 'twine,' found in Buli, Hausa, Katcha, Khoekhoe, Mbum, Ngambay, Rendille, Yoruba, Berik, Burarra, Gurindji, Kaluli, Kyaka, Lavukaleve, Muna, Rotokas, Kosarek Yale, Yir Yoront, Abzakh Adyghe (inter alia), Badaga, Basque, Bezhta, Chukchi, Khalkha, Kildin Saami, Sora, Cheyenne, Chickasaw, Ineseño Chumash, Haida, Kashaya, Lesser Antillean Creole French, Pawnee, Pipil, Quileute, Tuscarora, Wintu, Yuki, Central Yup'ik, San Lucas Quiaviní Zapotec, Copainalá Zoque, Arabela, Aymara, Bororo, Cashinahua, Cavineña, Hupda, Jarawara, Lengua, Piro, Tsafiki, Yanomámi, Bislama, Hani, Hawaiian, Bwe Karen, Manange, Samoan, Sedang, Takia, Tetun, and Yay (see Buck 1949: 550 for scattered Indo-European evidence). Somewhat similarly, Central Yup'ik has *qip'arpak* /qip'aq-rpak/ 'thick.hand-twisted.thread.large.' Many languages of Mesoamerica and adjacent areas colexify 'rope' with 'lasso.' This is the case in Itzaj, San Mateo del Mar Huave, Xicotepec de Juárez Totonac, Copainalá Zoque, and Cubeo. The relevant Xicotepec de Juárez Totonac term *tasiuj* is derived from *sihuí* 'to bend,' and the Wintu term contains a verb meaning 'to swing a rope or lasso.' In Rotokas, 'rope' is *iroiro*, which appears to be reduplicated from *iro* 'belt,' and similarly Bororo has *kogu-ia* 'belt-opening;' Kolyma Yukaghir colexifies 'belt' with 'rope.' Buli, Hausa (among other meanings), and Bislama colexify 'rope' with 'badge of rank' (Ngambay also with 'rank,' alongside 'trap'), Khoekhoe, Ngambay, Mali, Bezhta, and Cashinahua with 'fishing line' (Khoekhoe also with 'pore'), Carrier, Haida, Pawnee, Cubeo, Jarawara, and Hawaiian with 'line' more generally, Buli, Embera, and Hani with 'cable,' Pawnee and Hawaiian with 'thong,' and Kaluli, Chukchi, Cavineña, Jarawara, and Hawaiian with 'strap.' Muna *kalolai* 'rope for lowering things' is derived from *lolai* 'umbilical cord, to lower with a rope,' and Central Yup'ik colexifies 'rope' with 'umbilical cord.' Embera *hĩṅkará* also means 'whip, lash' when used with masculine gender and 'reed, cane' with neuter gender, and similarly, Hawaiian *kaula* also means 'whiplash,' alongside 'arc of circle' and "chain, as used by surveyors and engineers." Koyraboro Senni colexifies 'rope' with 'bundle of rice' inter alia, and Muna *kakoo* is derived from *koo* 'bundle of rice.'

Other associations include: Hausa *tuke*, meaning 'rope, thick string' in the dialect of Kano, otherwise means "twist together all of the material, e.g. as in making rope" inter alia. Ngambay *gól* also means 'to arrange, reconcile,' and Swahili *kamba* also 'shrimp, prawn.' Buin *kuuku* is also a female name. The Burarra term *murndurn* also means "group, work party, clan or tribe," and Kyaka *puu* also means 'bandage' among many other things, while *pungi* also means 'liver.' Ngaanytjarra *purturru* also denotes a 'hair string' (presumably used as a rope) and 'wool,' while One *apa* colexifies 'rope' with 'rattan.' Rotokas *koro-viri* seems to be analyzable as 'fruit-twist.something,' and Sko à also means 'clear, shiny.' Toaripi *horou* also means 'intestines,' Kosarek Yale *heing* also 'eye,' and Badaga *agga* ~ *hagga*

also ‘plow’ and ‘connection.’ The Basque term *soka* is also used with the meanings ‘halyard’ and ‘dress,’ and Khalka *ujayasu* also means ‘knot,’ while *degesy(n)* is also an obsolete unit of measurement. Sora *lua:dən* also denotes “fibrous bark” as well as a specific type of string “with knots to indicate payments made to creditors.” Chickasaw colexifies ‘rope’ with ‘yarn’ and ‘leash,’ and Lesser Antillean Creole French with ‘chord, note’ (presumably due to phonological collapse of French *corde* and *accord*). Kiliwa *t+ha?-q=h-?+nyat-tay-u* ‘lead rope’ is analyzable as ‘OBJ+mouth-ABL/ALL=3-??+pull-FREQ-PL,’ Nuuchahnulth colexifies ‘rope’ with ‘in line,’ whereas in the Santo Domingo de Guzmán dialect of Pipil, *kwerda(h)* is also ‘a measure of land.’ Tuscarora *uhsì'reh* also means ‘wick’ and ‘wire,’ Wintu *če'k* also ‘ropelike root(s),’ and *pəli* also ‘grapevine branches.’ Central Yup'ik *qecik* means ‘skin rind, scab’ and in the Norton Sound-Unaliq dialect also ‘rope.’ Copainalá Zoque colexifies ‘vein,’ and Bora *wáábya-u* is analyzable as ‘hammock-CL.round.’ Guaraní *sâ* also means ‘slavery,’ as well as, verbally, ‘to be fastened with a rope.’ Piro *tsa* also means ‘fiber,’ Wayampi *yā* also denotes a tree species, Fijian *dali* also ‘ten cuttle fish tied together,’ Bwe Karen *-bli* also ‘nest,’ Kapingamarangi *hali* also “to leak, to flow, to ooze,” Manange *Its'ho* also ‘swamp,’ and Bislama *rop* also ‘tape of cassette.’

#### 90. *The Scissors*

Representation: 71%

Motivated: 28.6%

Thereof Analyzable: 24.8%

Thereof Colexifying: 4.3%

Thereof by Contiguity: 19.3%

Thereof by Similarity: 7.9%

Recurrent associated meanings: cut, knife, tongs, cloth, razor, clip, pinch

Words for ‘scissors’ (or ‘shears’) are often derived from verbs meaning ‘to cut’ generally or more specific types of cutting, such as Blackfoot *sisóya'tsis* /*sisayi-a'tsis*/ ‘cut.into.strips-INSTR’ or Sedang *konep*, which is derived by the nominalizing infix <on> from *kep* ‘cut hair’ (see Buck 1949: 560 for the situation in Indo-European). Such terms are also attested in Chickasaw, Pawnee, Yaqui, Abipón, Ancash Quechua, Toba (where the term colexifies ‘knife’ and ‘razor,’ as is the case in Buli), Wayampi, Yanomámi, and Samoan, while there is a semianalyzable term in Bezhta. The Comanche, Kiowa, and Miskito terms feature an additional constituent meaning ‘cloth’ (e.g. Comanche *wana koo?* ‘cloth cutter,’ this term is archaic), and Muna colexifies the verbal reading ‘to cut’ (alongside “cross to the other side”) and the nominal one as ‘scissors’ directly. There are also languages which have a specific verb meaning ‘to cut with scissors’ (Biloxi, Arabela, Central Yup'ik), with the noun derived from it. Efik *ufań'kpö* contains *fat* ‘clip, cut with scissors’ and *ńk'pö* ‘thing,’ and similarly, Fijian has *i koti* ‘DERIV clip/shear.’ In two sampled languages, Japanese and Central Yup'ik (Nunivak island dialect), terms are derived from a verb meaning ‘to pinch’ rather than ‘to cut’: *hasam-i* ‘pinch-NR’ (colexifying ‘scissors of lobster’) and *nunuutek* /*nunur-(u)n*/ ‘pinch-device.for’ (this term is formally dual). Similarly, the Arabela term *tuquetaja* is derived by instrument nominalization (-*taja*) from *tuquenu* ‘to pince bare.’ A semianalyzable term involving a verb meaning ‘to pinch’ is found in Quileute. Otherwise, it is frequent cross-linguistically to have complex terms for ‘scissors’ based on other arti-

facts with similar function and an additional element indicating the *differentia specifica*, which is in this case often the presence of the holes serving as handles for the fingers. The source artifact is most commonly ‘knife’: Sko *tanglilong* contains *tang*, which is a general term for blades and hence also can refer to a ‘knife,’ and *long* ‘key, hole’ (there is a word *lí* meaning ‘cross-pole for roof or floor’ inter alia), Cheyenne has *hótâxová-mota* ‘crosscut-knife,’ and Central Yup’ik *nuussicuak* /*nuussig-cuar*/ ‘knife-little.one,’ the Yanomámi term for ‘scissors’ is that for ‘knife,’ amended by the quantal classifier *-ki* (see § 4.4.1), and there is a semianalyzable term featuring a constituent meaning ‘knife, sickle’ in Oneida, and a redundant term with a constituent meaning ‘knife’ is present in Mandarin Chinese (note also Tundra Yukaghir *ayñald’oyoje*, literally ‘knife with a mouth,’ Nikolaeva 2006: 106). Toba, as mentioned above, colexifies ‘scissors’ with ‘knife’ directly, and similarly, Buli *poning* (related to *poni* ‘to shave, cut hair’) also denotes a small knife for shaving. The cross-linguistic situation is thus exactly parallel to that within Indo-European: most frequent are terms derived from ‘cut,’ with the second most association being that with ‘knife’ (some Indo-European words for ‘scissors,’ notably in Celtic and Slavic, are formally the plural of or derived from words for ‘knife,’ Buck 1949: 560). Similarly, the handles are also used as the conceptualization source in Rotokas, which has *kaporoto*, containing *kaporo* ‘space between objects, (mountain) passage.’ Ket has a term making reference to both ‘knife’ as well as another perceptually similar artifact, ‘tongs’: *atəpəl do’n* /*atəp-ul do’n*/ ‘tongs-handle knife.’ An association with ‘tongs’ or ‘pliers’ is also realized by the complex Kapingamarangi term *di kabi-kabi* ‘ART fire.tongs-RED;’ *kabi* also has the additional meaning ‘to hold something between two other things.’ Furthermore, ‘scissors’ and ‘tongs’ are colexified in Yir Yoront, Khalkha, and Hawaiian (here, also with other implements similar in function).

Other associations include: Yir Yoront *thaminhwaw* contains *minh* ‘animal,’ and Badaga *katri ko:lu* contains *ko:lu*, meaning inter alia ‘stick’ and ‘skewer.’ Chukchi *wəŋətkunəŋ* contains *wəŋ* ‘yarn.’ Kildin Saami *rūvv’t* also means ‘iron’ and ‘trap.’ Haida has a term based on a verb meaning “make go (apart) into two pieces” prefixed with a verbal classifier for tongs or scissors. Nez Perce *capá·kakiwkaʔs* is analyzable as /*cepé·ké·kíw-ʔs*/ ‘by.pressure-with.teeth-take-??,’ and Nuuchahnulth *ʔap-ýak* as ‘straddle-tool.’ As a verb, Wintu *pín* also means ‘to squeeze’ as of long objects, as well as ‘scissors-like leg movements.’ Bora *majchówa* appears to be derived from *majcho* ‘food, eat’ (there is another term featuring an element meaning ‘house’ and ‘triangular frame’); perhaps there are errors in lemmatization in the consulted source. Cavineña *tishira* is also a term for wood that sustains the roof of houses (accidentally, if *tishira* < Span. *tijera*), and Hupda *hǎy’b’ah* is analyzable as ‘shear-flat.thing.’ Tehuelche *ʔepernwe* is derived from *ʔep’er* ~ *ʔep’ere* ‘to crop,’ and there is a semianalyzable term featuring a constituent with this meaning in Toba. Bislama *sisis* (< Engl. *scissors*) also means ‘close friend’ and, verbally, “to stick closely to someone, to hug sexual partner,” while Vietnamese *kéo* also means ‘to pull.’

91. *The Shoe*

Representation: 81%

Motivated: 29.4%

Thereof Analyzable: 20%

Thereof Colexifying: 9.4%

Thereof by Contiguity: 21.0%

Thereof by Similarity: 1.1%

Recurrent associated meanings: foot, put on/wear, sandal, footprint, moccasin, skin, walk

Most common among the lexico-semantic associations for ‘shoe’ (‘boot,’ ‘footwear’) is that with ‘foot’ (found also by derivation in Ancient Greek and by compounding in Welsh, Buck 1949: 428). Among terms betraying this association, a particularly frequent subtype is constituted by complex terms where the second constituent is a verb meaning ‘to put on’ or ‘wear,’ such as Kaluli *gib-a sa:ga:la:sen* ‘foot-?? put.on.’ Such terms are also found in Laz, Haida, Yuki, Miskito, and Bwe Karen; similarly, Samoan has *se’e-vae* ‘slide/slip-foot/leg.’ In Australia, it is particularly common to colexify ‘shoe’ with ‘foot’ directly. This is found in Gurindji, Ngaanyatjarra, Yir Yoront, dialectally in Basque (which otherwise has a derived term), as well as by a prefix in Comanche (Ngaanyatjarra, Yir Yoront, and Comanche also colexify ‘footprint,’ and Comanche also ‘trail,’ while Yir Yoront has an additional complex term containing elements meaning ‘human’ and ‘go in’). Otherwise, the secondary associations are manifold. Efik has *ik’pa-uküt* ‘leather-foot/leg,’ Ngambay *né-gòl* ‘thing-foot/footprint,’ Pawnee *asuuru*?, analyzable as /as-hur-u<sup>2</sup>/ ‘foot-place-NOM’ (originally denoting the ‘moccasin’ specifically, now ‘shoe’ generally;’ this type of autohyponymy is synchronically still present in Cheyenne, Nez Perce, and Yana, while in Buli, Hausa, Rendille, Sko, Toaripi, Hani, Hawaiian, Rotuman, and Samoan, relevant terms also mean ‘sandal’), Bora *túhapááji* /túhaá-pa:hi/ ‘foot-SCM.hole,’ Bororo *bure tadawu* ‘foot which.is.under’ (note that Greek and Sanskrit terms for ‘shoe’ is derived from a verb meaning ‘to bind under,’ ‘to tie under,’ Buck 1949: 428), Cashinahua *bin tae* ‘rubber foot,’ Guaraní *py-ao* ‘foot-clothes,’ Kaingang *pên né* ‘foot container,’ Maxakalí *pata-xax* ‘foot-cover/skin/bark’ (the Tuscarora term, non-transparent today, might have been made up of elements meaning ‘foot’ and ‘cover’ originally, and a compound featuring elements with these meanings is attested in Persian and is etymologically recoverable for Welsh, Buck 1949: 428), Yanomámi *mamikititioma* contains *mami* ‘foot’ and *titiha-i* ~ *titihi-ai* ‘put in,’ and Kapingamarangi has *hii wae* ‘package/wrap leg/foot.’ The Mali term *alēcharachi* is derived from *lēchar* ‘foot,’ and Chukchi *jeyət* is grammatically the plural of the word for ‘foot.’ Furthermore, Kyaka has *kimbu suu*, with *kimbu* meaning ‘foot, leg’ and *suu* being a loanword from Tok Pisin, that is, however, also a native lexical item meaning ‘drain, trench’ inter alia. Semianalyzable terms where the identifiable constituent is ‘foot’ or ‘foot, leg’ are found in Kemtuik, Sentani, Sko, Carrier, Upper Chehalis, Wappo, and Guaraní. Less frequent associations are that with ‘skin’ (which is one of the meanings colexified in one of the constituents of the Maxakalí term mentioned above), found in Meyah, which has *mek mei ofos* ‘pig coastal skin,’ and occurring in Ngambay by colexification, and that with ‘to walk’ in Abipón (*n-acaR-haR-late* ‘POSS.INDEF/3SG-walk-the.one.who-LOC;’ this term colexifies ‘ladder’). Badaga colexifies ‘walk on, step on’ and other meanings with ‘shoe.’ Furthermore, Yana *nikiiwau(na)* might contain the root *ni-*, meaning ‘for a male to walk.’

Other unique associations include: Hausa *takalmi* is also used metaphorically with the meaning ‘provisions for a journey,’ Muna *kolo* ‘wooden shoe’ (in this sense a loanword from Bahasa Indonesia ultimately going back to Dutch *klomp*) also indigenously means ‘sour’ and ‘to carry someone on the back,’ and Ngaanyatjarra *tjina* may, alongside ‘foot,’ ‘footprint,’ and ‘shoe,’ also refer to ‘claws, talons,’ ‘tracks’ and means ‘on foot’ adverbially. Rotokas *kuroea* is also the name of a species of vine with “leathery appearance,” and Basque *zapata* is also used with the meanings ‘threshold,’ ‘chassis,’ ‘frame,’ ‘buttress’ and “thin wooden fence.” Bezhta *halatco* is made up of the word for ‘leg’ and the essive case marker. The San Mateo del Mar Huave term *socol napiüc* appears to contain *socol* ‘corner,’ and Nez Perce colexifies ‘shoe’ with ‘horseshoe.’ Cubeo *curaido* consists of *curai* ‘ground’ and *-do*, the classifier for hole-like objects. Guaraní colexifies ‘shoe’ with ‘sock,’ Fijian *vāvā* also denotes the rungs of a ladder, Hani *seiqnaov* ‘shoes, sandals’ might be related to *seiq*, meaning inter alia ‘hoof’ and ‘bamboo stick on which to roll cotton in preparation for spinning it,’ Hawaiian *kā-ma’a* is analyzable as ‘CAUS-bind,’ Takia *su* also means ‘breast, udder, milk’ (due to collapse of an inherited term with Engl. *shoe*), and Yay colexifies ‘shoe’ with ‘to put out of the mouth.’

## 92. The Street

Representation: 89%

Motivated: 26.2%

Thereof Analyzable: 9.4%

Thereof Colexifying: 16.0%

Thereof by Contiguity: 9.9%

Thereof by Similarity: 6.2%

Recurrent associated meanings: manner/method/system, town/settlement, door, journey, go/walk, market, footprint, between, fare/freight, waterway, Milky Way

The most frequent association for this meaning (additional glosses such as ‘way,’ ‘trail’ etc. are disregarded in the following discussion) is a metaphorical abstraction, namely to ‘manner,’ ‘method,’ or ‘system.’ Eleven languages in the sample, Koyraboro Senni, Ngambay, Rendille (where the term also means “right thing to do,” “good way (of behaving)”), Kwoma, Basque, Greek, Guaraní, Rotuman, Samoan, Tetun, and Bislama feature this pattern of colexification. Similarly, the Burarra term is also used with reference to the ‘way of living,’ and the Kyaka term also means ‘category.’ In two areas of the world, New Guinea and the American Northwest, ‘street’ is colexified with ‘door’ and/or ‘doorway’ in some languages. This pattern is attested in Dadibi, Kaluli, Upper Chehalis, Haida, and Nuuchahnulth. Three sampled languages, Efik, Chickasaw, and Fijian have complex terms for street where one of the constituents has a meaning akin to ‘between,’ for instance Chickasaw has *okla-ittintakla* ‘town-between’ and Fijian *saqata ni koro* ‘interval POSS village.’ Indeed, associations with ‘town’ or ‘settlement’ are themselves relatively frequent. Badaga, Basque, and Ancash Quechua colexify these meanings (and Kildin Saami colexifies ‘street’ with ‘place in town’), Tuscarora has *yuta?nakáhrę?*, analyzable as /yu-ta’n- kahř(w)-?/ ‘3SG.NEUT.PATIENT-settlement-be.an.opening-STAT’ (compare Muna *kabhongka*, derived from *bhongka* ‘to crack, smash, for a road to open’). The relevant Buin term also is the name of a particular village. In four sampled languages, Chukchi, Abipón, Aguaruna, and Guaraní,

words for ‘street’ derived from verbs meaning ‘to go’ or ‘to walk’ are found (evidenced diachronically in Baltic, Buck 1949: 721); the relevant Guaraní term also means ‘pedestrian’ and ‘concubine.’ As an example of such a derived term, Chukchi *talan* /təle-n/ ‘go-LOC’ may serve; note also that the Bora term *ulléjito* is derived from *ulléje* ‘to travel’ by means of the classifier *-jito* for lines or roads. Moreover, Rotuman colexifies ‘street’ and ‘to go, come’ directly.

Khoekhoe and Piro colexify ‘street’ with ‘footprint,’ and similarly, Central Yup’ik has *tumyarak* /tuma°-yarak/ ‘footprint/track/trail-device.for.’ Buli, Ngambay, Basque, Greek, and Rotuman colexify ‘street’ with ‘journey,’ while Khalkha, Hani, and Mandarin colexify it with ‘market’ (and similarly, Yoruba with “public thoroughfare”), and Wayampi (alongside other meanings) and Hawaiian with ‘waterway’ (for which compare the association between ‘way, trail’ with ‘river’ reported in section 47). Similarly, Hausa colexifies ‘channel.’ Finally, Rotuman and Bislama colexify ‘street’ with ‘fare, freight,’ and Hausa and perhaps Tasmanian (Plomley 1976: 408) colexify ‘Milky Way’ (Hausa also ‘beehive’ inter alia).

Other associations include: Buli *siuk* also means ‘permission, right’ inter alia, and Efik *añ’wa*, denoting a ‘principal street,’ is analyzable as /a-ñwañ’a/ ‘NMLZ-widen’ (compare the derivation of a word for ‘street’ from ‘wide’ in Ancient Greek, Buck 1949: 720). Swahili *barabara* also means “proper, as it should be,” Buin colexifies ‘street’ with ‘access,’ and another Buin term, *rootu*, is the outcome of borrowing of both Tok Pisin *lotu*, meaning “church, worship, church service” and English *road*, and consequently has both readings. Kwoma *nobo* also means “illegitimate; born out of wedlock,” Lavukaleve *lake* also means ‘fire,’ and Muna *sala* is also used with the meanings ‘sort’ and ‘thing,’ alongside ‘trousers’ and other meanings. Toaripi *oti-haro* is analyzable as ‘place-head/chief,’ Rotokas colexifies ‘road, path, way’ with ‘line,’ and Yir Yoront *yalaq* also denotes groups of animate beings, such as a flock of animals or school of fish. Badaga *ke:ri* also denotes the ‘frontyard’ and “work-space in front of houses” and is “an appropriate measure of land area.” Basque *bide* also can refer to the ‘platform’ at a train station inter alia, *kale* also to the ‘eye of a needle,’ again next to other meanings. Greek *drómos* can also refer to a ‘distance,’ as well as ‘speed’ and a ‘race.’ Japanese *tōr-i* is analyzable as ‘pass-NR.’ Khalkha *yudumzi(n)* ~ *yudamzi* seems to be related to *yudum* “passage, thoroughfare, hallway, corridor,” *zegeli* is identical segmentally to one of the variants of *zegeli* ~ *zegele* ‘debt, loan,’ and Sora *taŋgo:rən* ~ *taŋo:rən* also means ‘occasion.’ Biloxi *nětkohi* ~ *natkohi* ~ *nětkohi* ~ *nětkuhi* ~ *nětkohi* contains *ně* ‘to stand,’ and *kohi* ~ *kuhi* ~ *kuhi* ~ *kuhi* ‘up, high.’ Comanche *kawonokatų* is said to literally mean “wolf separate camp.” Haida colexifies ‘street’ with ‘ladder,’ and Kashaya *hi?da* contains *?da* ‘extend, stretch.’ Aguaruna *jinta* also means ‘faculty, specialty’ and ‘terrain,’ Bora *huúva* might contain *huú* ‘tube,’ Bororo colexifies ‘railway’ and ‘trace of cobra,’ and Chayahuita *pa’térinso* is derived from *pa’térin* ‘to weave.’ Guaraní *tape* colexifies ‘street’ with ‘art’ and ‘religion,’ while the Miskito term *yabal* also means ‘mouth, tongue, way of speaking’ and *bila* also ‘mouth, opening,’ ‘center, inner part, space’ as well as ‘word’ and ‘language.’ Fijian *sala* also inter alia means ‘to climb or creep upon,’ and Bwe Karen *kle* also ‘to fix on’ inter alia. Kapingamarangi *ala* also means ‘responsibility’ inter alia, and Lenakel *suatu* also ‘course of ship’ and the “traditional exchange links between individuals or de-



scent groups." Malagasy *arabè* is analyzable as /*araka-bè*/ 'following-big/much.' Rotuman colexifies 'meaning,' 'condition,' and other things. Sedang *tróang* also means 'sentence,' Vietnamese *đường* also 'sugar,' and Yay *ran*<sup>1</sup> also 'to see.'

### 93. *The Table*

Representation: 77%

Motivated: 35.5%

Thereof Analyzable: 23.6%      Thereof Colexifying: 11.9%

Thereof by Contiguity: 19.5%      Thereof by Similarity: 7.9%

Recurrent associated meanings: eat, furniture, food, to place on, board, floor,  
flat, dish, tree/wood, place, restaurant

Terms for 'table' derived from verbs meaning 'to eat,' such as Carrier *ukwez-eʔal* 'on.it-eat' are found in Chickasaw, Kiliwa, Lake Miwok (where the term colexifies 'food, groceries'), Nez Perce (where the term colexifies 'restaurant, inn' and 'kitchen,' the former pattern is also attested in the Norton Sound - Upaliq dialect of Central Yup'ik), Wintu, Central Yup'ik (again, only in the Norton Sound - Upaliq dialect), Abipón, Arabela (where the term is derived by an agentive nominalizer and the resulting term hence can also refer to an 'eater, one who eats,' as is the case in Guaraní), Guaraní (where the term also means 'refectory,' 'dinner guest,' and 'tablecloth'), and Yanomámi. A semianalyzable term is present in Upper Chehalis, and similar complex terms of the lexical type are found in Biloxi, where *a'diʔoʔni ~ adiʔoʔ* contains *ti* 'to eat' and *oʔ ~ oʔni* 'to make, do,' Blackfoot (*iitáísooyoʔp* /iit-á-iso-ooyi-oʔp/ 'LOC-DUR-on.a.horizontal.surface-eat-21.NOM'), Cheyenne (*táxemésêhestôtse* /táxe-mésehe-hestôtse/ 'upon-eat-thing'), Haida (*ga taa daan* 'INDEF.PRONOUN eat place'), Kiowa (*p'í-á* 'eat-board'), and Yana (*moori?mauna* /ma-ri?mau(na)/ 'eat-place'). Moreover, Yir Yoront has *yo-penpn may pay+n* 'wood-flat food eat+NOUN.THEME.FORMATIVE' for 'dining table' specifically, and indeed, complex terms based on 'food' rather than 'eat' are also relatively frequent (and terms betraying this connection dominate in Indo-European, Buck 1949: 483). Thus, Kyaka has *nenge rate* 'food bench/shelf,' Kashaya *?aca? ?ama bumucid=tol* 'people food eat=on,' and Wintu *ba's top-i* 'food used.for-NOMINAL.STEM.FORMANT,' but a particularly frequent combination, especially in North America, is with verbs meaning 'to put,' 'to place on,' as in Oneida *atekhwahlákhwa?*, which is analyzable as /ate-khw-hel-hkw-aʔ/ 'SRFLX-food-set.on.top.of/place.on-INSTR-HAB.' Such terms are also featured in Tuscarora and Yuki, and, somewhat similarly, Pawnee has *rakaraaraaruukita?iitu?* /rakaraaraaruukita-iit-uʔ/ 'dishes-place-on.top.of-in.a.line-NOM,' Hawaiian *pā-kau* 'dish-put,' and *papa kau-kau* 'flat.surface RED-put' (though note that *papa*, which also has still other meanings, can also refer to a 'table' itself), Cashinahua *tsaun-ti* 'put-INSTR,' and Yuki also has the alternative term *həwəy ʔuʔk-ul* 'food dish-INSTR.' There is a semianalyzable term where the identifiable constituent means 'food' in Huambisa, and the association is present also in Abipón due to colexification of verbal 'eat' and nominal 'food.' In fact, in Lake Miwok and Yanomámi, which feature derived terms from 'eat,' 'table' is itself colexified directly with 'food.' In Cubeo, *tuoiwa* is derived from *tuoyu* 'to serve food' by means of the classifier *-va* for broad and flat objects, while Ket *laʔm* also itself means 'flat' as an adjective. Alongside

this and the Yir Yoront term already mentioned above, an association between ‘table’ and ‘flat surface,’ among other meanings, is found by colexification in Hawaiian and possibly in Samoan, where *laulau* is ‘table’ and *lau* both ‘leaf’ as well as a classifier for flat or thin objects. A similar association to that in Yir Yoront and Kiowa, where the relevant terms contain constituents meaning ‘wood’ and ‘board,’ is found in Bororo, which has *irá* /i-ra/ ‘tree-bone.’ Baruya, Dadibi, Toaripi, and Hawaiian colexify ‘table’ with ‘floor,’ and finally, it is not uncommon for ‘table’ to be colexified with other pieces of furniture, which happens in Noni, Baruya, Burarra, One, Rotokas (by the term *koara ua*, putatively analyzable as ‘put.together NARROW.OBJECT’), Toaripi, Basque, Ket, Khalkha, Welsh, Wintu, Bororo, Cashinahua, Embera (also with ‘saddle’), Hawaiian, and White Hmong, in which case the common semantic denominator seems to be that pieces of furniture are typically human-made raised surfaces. Similarly, Ket, Welsh, Ineseño Chumash, and Bororo colexify ‘table’ (also) with ‘board.’

Other associations are: One *simpa* also means ‘bridge,’ and Khalkha *sirege(n)* colexifies ‘table’ with ‘throne’ as well as ‘altar’ and ‘feast, banquet.’ Welsh *tabl* also means ‘tablet.’ Nez Perce *tí’mewe’s* is analyzable as /tí’mē-nwe’s/ ‘write-LOC,’ and Wintu *panti?ilestopi* ‘table, desk’ contains *panti* ‘on top, on’ and *top* ‘used for.’ Guaraní *arikapa* is a neologism containing *ári* ‘on top’ and *pa* ‘everything.’ Hani *lolbieil* contains *bieil*, meaning ‘wide’ and also acting as a classifier for ‘chairs’ and ‘tables.’ Hawaiian *papa* means ‘flat surface’ generally, and therefore colexifies ‘table’ with notions such as ‘reef,’ ‘layer,’ and others, while *pūne’e*, meaning also ‘movable couch’ and rarely ‘pew,’ appears to be related to *ne’e*, which means ‘moving along little by little or by fits and starts, to step’ inter alia. The Sedang term *ko’bang* also means ‘blackboard,’ and Yay *soŋ*<sup>4</sup> means ‘to scrape’ in verbal usage.

#### 94. *The Toilet*

Representation: 50%

Motivated: 54.4%

Thereof Analyzable: 49.3%

Thereof Colexifying: 6.8%

Thereof by Contiguity: 46.1%

Thereof by Similarity: 2.4%

Recurrent associated meanings: faeces, house, bathroom, place, bush/forest, hole, small, bathe/bath, outside, room, relieve, water, wash

The most common association for ‘toilet’ (or ‘latrine’) is by functional contiguity with ‘faeces’ or ‘defecate.’ Within terms betraying that association, the most dominant subpattern are terms of the lexical type with ‘house’ being the additional constituent, as in Dongolese Nubian *úññĩkã* /úññĩ-ŋ-kã/ ‘excrement-GEN-house.’ Such terms are also found in Angkor, Baruya, Dadibi, Kaluli, Kwoma, Kyaka (with an additional constituent meaning ‘origin’), Rotokas, Comanche, Kashaya, Santiago Mexquititlan Otomí, Pawnee (with an additional constituent meaning ‘to be going’), Wappo (which also has *mahkúyema* *chùya*, containing an element meaning ‘exit’ alongside ‘house’), and Yay, some of them denoting an ‘outdoor toilet’ specifically. Note also that the Bora term *naméja* is derived from *name* ‘excrement’ by the classifier *-ja* for houses, as well as Kiowa *sh’tsoue-tou* ‘urine-house’

for 'urinal.' Otherwise, 'place' is the meaning of the second constituent in Kwoma *she pii eem* 'faeces defecate place,' Japanese *ben-jo* 'faeces-place,' and Ket *horoks ba'ŋ* /ho'q-d-oks ba'ŋ/ 'excrement-POSS-stick place,' Hawaiian *wahi ho'o-pau pilikia* 'place CAUS-finish trouble' (which also has *ki'o-na* 'excrete-NMLZ,' colexifying 'dung heap, dump'), while Greek directly colexifies 'place' and 'toilet' (as well as 'role' and 'party'). Other terms of the lexical type with 'faeces' are Ngambay *buwá-sin* 'hole-faeces' (compare Kyaka *yuuwuali anda ~ iwali anda* 'deep.narrow.hole/pit house,' and the colexification of 'hole' with 'toilet' in Hawaiian, among other meanings), Hupda *y'ǎ-ʔteg* 'faeces-thing,' and Hani *xiqduq*, with *xiq* meaning 'faeces' and *duq* meaning 'to dig,' but also functioning as a classifier for pools of water and the sound of drums or thunder. Derived terms also occur, namely in Central Yup'ik, Aguaruna, Bora (as mentioned above), Wichí, and perhaps in the Cuisnahuat dialect of Pipil; semianalyzable terms are found in Chukchi, Nez Perce, and Guaraní. Similarly, Yoruba *ibi-igbònsẹ* is analyzable as 'place-relieve,' and Gurindji has *warlp-kaji* 'relieve.self-AGENT.'

Complex terms where one of the constituents means 'house' are also found with second constituents other than 'faeces:' Bislama has *smol-haos* 'small-house,' which is paralleled in Khoekhoe and Samoan, Abzakh Adyghe has *psə-wəne* 'water-house' (for the association with water, compare also Koyraboro Senni *hari-mun-doo* 'water-pour-place'), Welsh *tŷ bach* 'corner/hook house,' Cheyenne *o'ěhné-mâhéó'o* 'eliminate-house' and *too'hamé-mâheo'o* 'bathe house,' Santiago Mexquititlan Otomí analogously *nguu nsaha* 'house bath' (Carib moreover has a derived term from 'bath'), Samoan *fale-'ese* 'house-other,' *fale-ui* 'house-go.along,' and *fale-vao* 'house-bush/forest' (note also the colexification of 'toilet' with 'bush/weed' or 'forest' in Ngambay and Lenakel), and Yay *raan<sup>4</sup> pay<sup>1</sup> rok<sup>5</sup>* 'house go outside' for an 'outdoor toilet' specifically (note also the colexification of 'toilet' with 'outside' in Aguaruna); further, Hausa *ban'daki* is a contraction of *bayan daki* 'back house.' Semianalyzable terms where one of the constituents is 'house' are found in Rotuman and Samoan. Japanese *keshō-shitsu* is questionably analyzable as 'mascara-room,' and Vietnamese *phòng vệ sinh* as 'room protect life.'

Moreover, terms involving a verb meaning 'to wash' are found in Japanese and Bororo. Finally, 'toilet' and 'bathroom' more generally are colexified in Buli, Khoekhoe, Swahili (alongside 'prayer mat'), Cheyenne, Kashaya, Aguaruna, Bororo, Hawaiian, and Tetun.

Other associations include: Hausa colexifies 'toilet' and 'cesspit,' and another term also denotes the "[d]oubling of a consonant, and the sign over a written consonant to indicate the doubling" as well as "a kind of European-made silk material." Swahili *choo* is related to *oga* 'to wash body.' Chickasaw *aaombiniili' ~ aambiniili'* is analyzable as /aa-on-biniili-'/ 'LOC-APPL-sit.down-NMLZ' and also means 'chair.' San Lucas Quiaviní Zapotec *ta's x:tèe'n ba'nny* is analyzable as 'bowl of bathroom,' and Toba *nogoxoqui* is related to *alogo* 'clothes.' Rotuman *fā'u* inter alia also means 'back' (note also that Buli *banjiri* borrowed from Hausa *bayan gida* 'behind the house').

95. *The Train*

Representation: 41%

Motivated: 52.2%

Thereof Analyzable: 45.6%

Thereof Colexifying: 6.7%

Thereof by Contiguity: 10.0%

Thereof by Similarity: 38.9%

Recurrent associated meanings: wagon/vehicle/car, fire, iron/metal/steel,  
boat/ship, smoke, land/ground, run, road/path, carry

It is common for motivated terms for ‘train’ to be of the lexical type, with one of the constituents meaning ‘wagon,’ ‘car,’ or ‘vehicle’ generally. The semantics of the second constituent varies considerably, but not without limits. Terms with ‘fire’ or ‘fiery’ being the second constituent are found in Abzakh Adyghe (*meś<sup>o</sup>ek* /maś<sup>o</sup>e-k<sup>o</sup>ə/ ‘fire wagon’), Khalkha, Kashaya, Lake Miwok, Hawaiian, Mandarin, and Vietnamese; note also Nez Perce *ʔalahín* /ʔá·la·hí·n/ ‘fire-thing.with’ and Yay *rua<sup>i</sup> fí* ‘boat fire.’ There are also languages where ‘smoke’ or ‘steam’ is the meaning of the second constituent, as in Swahili *gari* (*la*) *moshi* ‘car (of) smoke,’ and also in Japanese (for a ‘steam locomotive’ specifically), Blackfoot and Chayahuita; note also that Wichí *wapulh* is a loanword from Spanish *vapor* ‘steam.’ Kanuri has *màrà cídí-bè* ‘vehicle ground-of.’ Two Oceanic languages, spoken on small islands, showcase, similar to the situation in Yay just mentioned, a metaphorical transfer from ‘ship’ to ‘train,’ with the differentia specifica indicated by a term for ‘land’ or ‘ground’: Rotuman *ahai la’ ufa* contains *ahai* ‘ship’ and *ufa* ‘land,’ and Fijian *sitima ni vanua* is analyzable as ‘steamer POSS land,’ while Hausa colexifies ‘train’ with ‘boat’ (and ‘through,’ see section 76 on this pattern). White Hmong has a complex term where ‘iron’ is one of the other constituents alongside ‘wagon’: *tsheb nqaj hlau* ‘vehicle rail iron;’ other complex terms on the basis of ‘iron’ are Cheyenne *ma’aataameo’o* /ma’aata-meo’o/ ‘iron-road,’ Yuki *lil haʔ-ol ~ lil ham-ol* ‘iron carry-INSTR’ (compare Kiliwa *waʔ=t-kw+lkwi-y-tay* as ‘house=OBJ-WH+carry-ATT-FREQ’ with the literal meaning “back-packer house” according to the source), and Malagasy *fiarandalamb’y* /fiara-n-làlana-vy/ ‘vehicle-GEN-road-iron.’ Santiago Mexquititlan Otomí colexifies ‘iron’ with ‘car,’ and ‘train’ is *ma-bojá* ‘long-iron/car’ or *njunubojä* /n-juni-bojá/ ‘NMLZ-to.tie-iron/car.’ Biloxi has *yaduxta<sup>n</sup> ta<sup>n</sup>hi<sup>n</sup>* ‘wagon run’ and a verb meaning ‘to run’ also figures in the Chickasaw term (*itti*) *chanaa malili-* ‘wood roll run-NMLZ.’ Other complex terms with ‘vehicle,’ ‘car,’ or ‘wagon’ are Khoekhoe *fnū-kuni-s* ‘black-wagon-3SG.FEM,’ Japanese *den-sha* ‘electric-car,’ and Oneida *tsyoʔslehtá’kat*, analyzable as /s-yo-ʔsleht-ʔkaht/ ‘REPETITIVE-NEUT.PATIENT-vehicle-move.fast.’ Moreover, a semianalyzable term with ‘wagon’ is found in Blackfoot and with ‘bullock cart’ in Bwe Karen, while Comanche, Pawnee, and Wintu colexify ‘train’ with ‘car’ directly.

Other associations include: Buli *girigiri* is also an onomatopoeic word that “imitates a low rumbling noise” generally. Khoekhoe *lnubutas* contains the verb *lnubu* ‘to shake, agitate,’ and Gurindji *turrkalangarna* contains *ngarna* ‘denizen.’ The morphological analysis of Carrier *yenkenekhēs lēnedīzṭī* remains unclear, but the “literal” translation offered by the source for *yenkenekhēs* is “land on it (being heavy) moves on,” and *lēnedīzṭī* contains elements meaning “attached together” (*lē*), “in a line” (*ṭī*) and “being several” (*ne*). Kiowa *ʔnk’iṭ-gā* is analyzable as ‘go.along-NOUN.POSTFIX,’ while Lesser Antillean Creole

French *twen* (< Engl. *tren* 'train') also means 'attendants' as well as 'noise, trouble.' Samoan *nofoa afi* is analyzable as 'sit-SUFFIX (fire/engine)' and colexifies 'chair, seat' and 'saddle.'

#### 96. *The Weapon*

Representation: 51%

Motivated: 50.9%

Thereof Analyzable: 38.7%

Thereof Colexifying: 14.2%

Thereof by Contiguity: 34.7%

Thereof by Similarity: 0.9%

Recurrent associated meanings: tool, thing, rifle/gun, war, fight, sharp, army,  
bow, wound/to wound, defend, hurt, kill, throw, soldier, use, battle

Terms for 'weapon' are sometimes derived from verbs meaning 'to fight,' as in Xicotepec de Juárez Totonac *lita'lān* /lī-ta'lá-n/ 'INSTR-fight-INSTR' (and also in Sanskrit, Buck 1949: 1384). Such terms are also found in Kashaya, Miskito, and Malagasy, while Rotokas *upo purapara* contains *upo* "strike, slap, fight, murder" and *pura* "make, do, create," Yoruba has *ohun-jà* 'thing-fight,' Samoan *mea-tau* 'thing-fight,' and there is a semianalyzable term in Blackfoot. There are also terms based on verbs with similar but different semantics. Chickasaw has *ishhottopachi* /isht hottopa-chi-/ 'with hurt-CAUS-NMLZ,' Hawaiian *mea hō'eha* 'thing hurt,' and a semianalyzable term containing a verb meaning 'to aim, hurt' is found in Yanomámi. The Carib term *owotopo* is derived from *owo* 'kill,' and Hawaiian has *mea pepehi kanaka* 'thing beat/kill man.' Laz *o-tōç-aše* is analyzable as 'DERIV-throw/shoot-DERIV' (similar derivation is found in Sanskrit for a term probably originally denoting missile weapons, Buck 1949: 1384), and Hupda *d'ap-b'ŷy-teg* perhaps as 'flesh-throw-thing.' Hausa colexifies 'weapon' with 'wound inflicted by a metal weapon,' and Nuuchahnulth *saʔa-čakʷ* is analyzable as 'wound/strike.with.weapon-tool.' More frequent than any of these associations, however, are terms where one of the constituents is 'war,' as in Kildin Saami *tuarr-viešš* 'war-thing.' Precisely parallel terms are found in Efik, Katcha, Noni, Yoruba (where 'war' is colexified with 'army' and 'battle' inter alia), Hawaiian (where the same remarks as for Yoruba apply), and Takia. Similarly, Carrier has *pê-netsepah* 'where-with-to.war' (alongside *pê-neznî'u* 'wherewith-to.be.cuirassed'). A literal translation of Bora *ñneri metéhmémeine* is "that with which we guard (defend) ourselves," and a Toba term for 'weapon' is derived from a verb meaning 'to defend.' A frequent pattern of colexification is that with (metal) 'tool,' found in Yoruba (by the analyzable term *ohun-i-lo* 'thing-NMLZ-use,' for which compare Tuscarora *yéčthaʔ*, which is based on the root -ačT- 'to use'), Badaga (where the relevant term also means 'door-bolt,' among other specific tools), Khalkha, Sora, Welsh, Highland Chontal, Lesser Antillean Creole French, and Central Yup'ik; this pattern is also common in Indo-European (Buck 1949: 1383-1384). Analogously, Japanese has *hei-ki* 'soldier-tool' (for the association with 'soldier,' note also Khalkha *cerig yn zemseg* containing *cerig* 'warrior, soldier, army' and *zemseg* 'ornaments,' Mandarin *wu3-qí4* 'military/valiant-utensil/appliance/apparatus,' and Vietnamese *vũ khí* 'martial tool.' Similarly, Guaraní *rairô rembiporu* is analyzable as /rairo tembiporu/ 'very.sharp instrument,' and San Mateo del Mar Huave colexifies 'weapon' with 'iron' and 'prison.' Similarly to the situation in Guaraní, Toaripi has *mare etau* 'sharp thing,' and Cheyenne *heškováñö'o*

contains *heškóv* ‘sharp’ and *ôhëö’o* ‘thing,’ the tentative literal translation offered is “sharp-by.hand(thing)” and is marked as unsure in the source. Still more frequent, however, is colexification with a particular type of ‘weapon’ (autohyponymy). This is found with ‘rifle’ or ‘gun’ in Greek, Highland Chontal, San Mateo del Mar Huave, Xicotepec de Juárez Totonac, Yaqui, and Miskito (there are also two languages with complex terms on the basis of ‘rifle’: Chayahuita has *irapa inapita pochin ninin-so* ‘rifle these like do/be-3SG.SUB’ and White Hmong the dvandva compound *riam phom* ‘knife gun’), with ‘bow’ in Kyaka and Bororo, with ‘sword’ in Badaga, with ‘slingshot’ in Cheyenne, with ‘machete’ and, by a different term, with ‘lance,’ in Chayahuita.

Other associations include: Hausa *makami* is derived from *kama* ‘to catch,’ and indeed can also refer to a ‘catcher.’ The Muna term *ewanga* ~ *iwanga* contains *ewa* ‘martial arts, to attack, oppose,’ while *sandata* also means ‘to supply, prepare’ and ‘supplies’ (the meaning ‘weapon’ is due to borrowing of Bahasa Indonesia *senjata*). Welsh *erfyn* also means ‘to beg, to pray,’ and the Abzakh Adyghe terms contains the word for ‘hand,’ as does the Haida term. Kiowa *mòn-pà-tò* is analyzable as ‘hand-against-hold.’ Bora *lliñája* ‘hunting weapons’ might contain *lliñe* ‘punishment’ and may indeed also refer to ‘hunting,’ there is a further derived term containing an element with that meaning. Santiago Mexquititlan Otomí *bojá* also means ‘money,’ Bororo *boe eiga* appears to contain *boe* ‘thing’ and perhaps *iga* ‘splint.’ Cayapa colexifies ‘shadow, soul, reflection’ (presumably due to collapse of Span. *alma* ‘soul’ and *arma* ‘weapon’), Guaraní *pojoapy* contains *apy* ‘deposit, extreme point,’ Huambisa *manitai mani* ‘warrior,’ and Jarawara *tahi/tahi* also means ‘killer,’ ‘hunter,’ ‘sliver,’ and is also the name of a song about a particular spirit. Piro has a semianalyzable term with one constituent meaning ‘man, person,’ and Toba *n’ataxaqui* colexifies ‘ammunition pocket’ and ‘spell.’ Yanomámi *shëmotima* contains *shë* ‘to hit,’ and Hawaiian *mea make*, also meaning ‘corpse,’ is analyzable as ‘thing die.’

#### 97. The Window

Representation: 71%

Motivated: 44.6%

Thereof Analyzable: 36.9%

Thereof Colexifying: 7.7%

Thereof by Contiguity: 19.5%

Thereof by Similarity: 16.0%

Recurrent associated meanings: door, hole/opening, light/bright, see, house, mirror, glass, small, eye, wind/breeze, entrance

Terms for window are derived from verbs meaning ‘to see,’ as in Cubeo *jáicobe*, presumably analyzable as /jáino-cobe/ ‘sight-CLASS.HOLE.LIKE.OBJECT,’ in seven sampled languages, alongside Cubeo also in Muna, Sora, Upper Chehalis, Comanche, Kiowa, and Arabela, where the derivation mechanism is agentive nominalization and the resulting term hence can also refer to ‘one who sees.’ Furthermore, terms with ‘see’ of the lexical type are found in Khoekhoe (*mû-tui-dao-s* ‘see-peep.through.opening-door-3SG.FEM’), Kiliwa (?+wa?=t-p+sa-w-tay-u? ‘DN+house=SUBJ-MP+see-DIR-FREQ-OBL’), and Pawnee (*uka’aatawiiriku* ~ *ukaatawiiriku* /*uka’aata-wi-iirik-hus*/ ‘shadow-??-see-IPFV’). The latter term colexifies ‘window’ with ‘glass’ in general and with ‘mirror’ in particular. Colexification with ‘glass’ is also found

(sometimes by analyzable terms) in Upper Chehalis, Oneida, Fijian (by the analyzable term *i ilo-ilo* 'DERIV look.at.reflection-RED), and Tuscarora; Nuuchahnulth has *pipihiqsim*, which is reduplicated from *pihiqsim* 'glass, pane of glass,' Hawaiian has *puka-aniani* 'opening/door-glass/mirror/transparent,' and Yay *tu<sup>1</sup> ciay<sup>2</sup>* 'door mirror/glass.' Colexification with 'mirror' (again, sometimes by analyzable terms) is attested in Upper Chehalis, Kashaya, Kiowa, Pawnee, Tuscarora, Wintu (here the relevant term contains elements meaning 'in, be in, put in' inter alia and 'male person' which is related to a more general reconstructed root 'person'), and Fijian (by the analyzable term mentioned above). Associations with 'door,' as seen in the Yay term just mentioned, is, however, more frequent. It is realized by colexification in Itzaj, Cashinahua, Bororo, Jarawara, Hupda, Miskito, Fijian, Hawaiian, and Sedang, which colexifies also other meanings. Morphologically complex terms with 'door' are exclusively of the lexical type in the languages of the sample. There are a variety of subtypes. Commonly, 'small' or 'little' is the meaning of the second constituent in such terms, as in Toaripi *utape seika* 'door small,' and also in Biloxi, Kiliwa (where a third constituent 'house' is also featured), Abipón, Bororo, and Malagasy. Some of these terms can also be interpreted compositionally and hence also refer to a small door. Note also Chickasaw *okkisoshi* / *okkisa'-oshi* / 'door-son' (and that the Ancient Greek word for 'window' is a diminutive of that for 'door,' Buck 1949: 470). Furthermore, Rotokas has *kiuvu ratao* / *kiuvu rataoa* / 'wind door,' for which compare also Kanuri *kàsàm-rám* 'breeze-place.of,' Yoruba *ojú-aféfé* 'eye-wind,' which precisely mirrors the etymology of English *window*, and Hawaiian *puka makani* 'opening wind' (this term also means 'opening for ventilation' and 'anus;' compare also the connection of Spanish *ventana* with Latin *ventus* 'wind,' Buck 1949: 470). Khoekhoe has *mû-tui-dao-s* 'see-peep.through.opening-door-3SG.FEM,' Abzakh Adyghe *šen(ə)-y°ə-pše* 'head/upper.part-RELAT/EPEN-??/end-door' (lexicographers note that the term is non-transparent nowadays), and Rotuman *nu'suar mutu* 'door cut.across/sever/cut.off.' In Greek, *paráthuro* is derived from *thúra* 'door,' and a semianalyzable term containing a morpheme meaning 'door' is featured in Yuki. Some of the terms colexifying 'window' with 'door' are morphologically complex, having constituents meaning 'house' and 'opening' or 'hole.' This is true of Itzaj (*jol-naj* 'hole-house'), Tuscarora (*yunęhsáhrarę* / *yu-nęhs-ahrar-ę* / '3SG.NEUT.PATIENT-house-be.a.hole-STAT'), Bororo (*wai-poro* 'house-opening'), and Hawaiian (*puka hale* 'opening house,' which also means 'door'). Kwoma has a term with an identical structure (*aka siisiwey* 'house hole'), but uses it only for 'window,' not for 'door.' Associations between 'window' and 'hole, opening' are also found in other configurations, however (just as in Indo-European languages, Buck 1949: 470). Koyraboro Senni has *fun-tarey* 'pierce.hole-area.outside.' This term is at the same time a loanword from French *fenêtre* that is "vaguely intelligible" as a Koyraboro Senni compound according to the consulted source (that is, presumably, it has been folk-etymologized), Kolyma Yukaghir *pońqā-šeščā* 'light entrance' and *pońqā-sōjnubəd-anil* 'light-enter-hole' (recall also that the means of derivation in the Cubeo term mentioned earlier is a classifier for hole-like objects, and note also that Hupda *məyó* is etymologically probably *məy-nə* 'house-mouth'), and Nuuchahnulth *kuukuuhsim* contains *kuh<sup>w</sup>* 'opening' and the lexical suffix *-sim* 'at an opening.' Furthermore, Kaingang has *kanē nor* 'eye opening,' Burarra, Cashinahua, Manange, and Sedang directly colexify 'hole, opening' with 'window,' and the Buli term also denotes

other kinds of openings in human-built structures. A semianalyzable term with this structure is furthermore found in Yay, and in Hani, *caqbyu* contains the classifier for holes, *byu*, with *caq* glossed as ‘to guess, conjure up,’ ‘to have something break’ inter alia. There are also terms in other languages which, as in Kolyma Yukaghir, betray an association with ‘light’ or ‘brightness’: Baruya *yiritinna* /yirite-tinna/ ‘daylight-eye,’ Chukchi *qeryās-ʔən* ‘light-one.who.has,’ Cheyenne *vó’nâhenhestôtse*, containing *vó’n* ‘light’ and *hestôtse* ‘thing,’ Chickasaw *aahashtahli* ~ *aahashtaali* /aa-hashtahli-/ ‘LOC-be.bright-NMLZ,’ Haida *radagats’aa7u* /radagaa-ts’a-7u/ ‘be.daylight-inside-INSTR,’ Lakhota *ožǵǵaglepi*, the literal meaning of which is ‘light frame’ according to the consulted source, Central Yup’ik (Nuni-vak island dialect) *tanqiun* /tanqik-(u)n/ ‘brightness-device.for’ for a “seal-gut skylight window,” and Miskito *ingni dimaika* ‘light entrance’ are of this kind (see Buck 1949: 470 for similar Germanic evidence; for the association with ‘day,’ note also that there is a semianalyzable term with a constituent meaning ‘day’ in Blackfoot). Moreover, Santiago Mexquititlan Otomí *neki xi hño* contains *neki* ‘seem’ and *hño* ‘good’ and also means ‘clear, bright.’ As seen in the Yoruba and Baruya terms already mentioned, metaphorical transfer from ‘eye’ to ‘window’ is another recurrent association. It is also attested in Mbum (*nják yâr* ‘verge/entrance eye’ and *yâr pàk* ‘eye house’) and Kaingang (*kanẽ nor* ‘eye opening’), and there is a semianalyzable term in Jarawara (see Buck 1949: 470 for further evidence for this connection in Germanic, Slavic and Indo-Aryan). As has also become clear already from the discussion so far, ‘house’ is a frequent constituent in terms for ‘window’ as well. Alongside the terms already mentioned, terms containing an element meaning ‘house’ are also found in Carrier (*tadĩntaz*, containing *ta* ‘house’ and *ʔaz* ‘gashing, slashing’) and Bora (*jaatu páhoowa nééne* /jaá-tu páhoo-wa néé-ne/ ‘house-ABL complete-CL.empty.space seem-CL.inan’),

Other associations include: Hausa *taga* also means ‘to begin, attempt,’ and *sagata*, which means ‘crossbeam’ inter alia, assumes the meaning ‘window’ in the dialects of Sokoto and Zaria. Dongolese Nubian *šibbāḡ* also means ‘lattice,’ and the Mali term *manaingĩēmgi* also means ‘torch, flashlight.’ Ngaanyatjarra *winta* is also used with the meaning ‘glasses,’ while Badaga *halagaŋŋi* also means “lookout, watchtower, observation point.” In Ket *qoqpul* “small window made of ice in the winter dugout,” the sequence *qo* may go back etymologically to *qō* ‘ice.’ Oneida *o’wíse?* also means ‘ice’ (and ‘glass’). Central Yup’ik *egaleq* ~ *legaleq* contains *ega-* ‘to cook by boiling,’ an association that is natural given that the term has been extended to ‘window’ from its original meaning “smokehole of a traditional Yup’ik house.” Cayapa *juucapa* appears to contain *capa* ‘point,’ Fijian has the obsolescing term *i kasivi bale* ‘DERIV to.spit drop’ (with *i kasivi* meaning ‘spittle’), and Samoan (*’o le*) *fa’a-malama* is analyzable as ‘REL ART CAUS-blaze.up.’ Finally, the Yay term *paak<sup>2</sup> taan<sup>2</sup>* is analyzable as ‘mouth different.’

#### 98. The Adam’s Apple

Representation: 47%

Motivated: 53.4%

Thereof Analyzable: 37.0%

Thereof Colexifying: 20.0%

Thereof by Contiguity: 22.9%

Thereof by Similarity: 28.5%



Recurrent associated meanings: throat/neck, larynx, windpipe, apple, bone, protrude/protuberance, fruit, trachea, goiter, knot

By contiguity, the most frequent association for the ‘Adam’s apple’ is that with either ‘throat’ or ‘(front part of) neck.’ Buli, Burarra, Dadibi, Kaluli, Muna, Wintu, Yuki, Fijian, and Hawaiian directly colexify ‘Adam’s Apple’ with one of these meanings. There are many semianalyzable terms featuring one of the abovementioned meanings (namely in Bakueri, Rendille, Biloxi, Upper Chehalis, Comanche, and Macagúan). Fully analyzable complex terms with ‘throat’ include Khoekhoe *dom-lkhom-s* ‘throat-bundle-3SG.FEM’ (Khoekhoe also has *domo-ro-s* ‘throat-DIM-3SG.FEM’), Yoruba *kókó-òfun* ‘wen/knob/excrescence-throat,’ Comanche *pia kuitsi?* /*pia kuitsi*/ ‘big throat,’ Toaripi *kōvōre luka* ‘neck.front/larynx stomach,’ Haida *qagan skuji* ‘throat bone,’ Lesser Antillean Creole French *zo gòj* ‘bone throat’ (compare also Wappo *lèkiš citi-wélma*, which is perhaps analyzable as ‘swallow bone-protection’), and Pawnee *raruucpaca*, analyzable as /*raruus-paca*/ ‘throat bump,’ while Guaraní colexifies ‘throat’ and ‘Adam’s Apple’ by the analyzable term *jyryvikandu* /*jyryvykandu*/ ‘throat-protuberance.’ Similarly, Samoan has *pona-ua* ‘knot/node/protuberance-neck,’ Ineseño Chumash *šoyokonič* is derived from *oqyokonič* ‘be uneven, protruding’ and Hawaiian colexifies ‘Adam’s Apple’ with “[a]ny kind of protuberance from a pimple ... to a hill.” The association with ‘knot’ is also found in another Austronesian language, namely Tetun, which has *kaka-fukun* ‘throat-knot/joint/knuckle.’ Associations with ‘apple’ in this configuration are found in Basque (*zintzur-sagar* ‘throat-apple’) and Itzaj (*mansaanaj-il kal* ‘apple-poss neck/throat’). Given that Itzaj *mansaanaj* < Span. *manzana*, as well as the fact that complex terms with ‘apple’ are found also in the other sampled European languages, namely Greek (*mī-on toú Adám* ‘apple-NOM.SG.NEUT ART.DEF.GEN.SG.MASC Adam’) and Welsh (*afal breuant* ‘apple windpipe’), this pattern seems to be originally peculiar to Europe, see § 6.4.3.3. for historical evidence corroborating this conjecture.

Complex terms with ‘neck’ include Toaripi *kōvōre luka* ‘neck.front/larynx stomach,’ Yir Yoront *man-nhapn* ‘neck-egg’ and *man-pung* ‘neck-sun,’ Wintu *dokikenti*, perhaps analyzable as /*dok-i-kenti*/ ‘neck-??-under,’ Arabela *riquia-nu* ‘neck-CLASS.PATH,’ Cayapa *cutu pijpuca* ‘neck wheel,’ Chayahuita *cono-pira* ~ *conu-huiru* ‘neck-pile,’ Guaraní *ajukytâ* /*ajura-kytâ*/ ‘neck-wart’ (alongside *ñoko’ê-kytâ* ‘throat-wart’), Fijian *lobi ni domo* ‘fold.lengthwise.or.crosswise POSS neck/voice/sound,’ and Hawaiian *kani-ā’ī* ‘hard-neck’ (which however, can refer to the ‘neck’ itself inter alia; another Hawaiian term, *pū’ā’ī*, probably goes back to \**pu’u* ‘ā’ī ‘protuberance neck’). Semianalyzable terms are found in One, Yir Yoront, Nivkh, Yana, Toba, and Yay.

Lavukaleve colexifies ‘Adam’s apple’ with ‘fruit,’ Sedang has *plai plê* ‘fruit/seed/pellet armspan,’ and a semianalyzable term with ‘fruit’ is attested in Lenakel. Khoekhoe, Dongolese Nubian, Burarra, Muna, Badaga, Khalkha, Aguaruna, Arabela, Carib, Cashinahua, Fijian, and Hani colexify ‘Adam’s Apple’ with ‘larynx’ (Upper Chehalis has a semianalyzable term with this being the meaning of the identifiable constituent), Kyaka, Wintu, and Carib with ‘windpipe,’ and Chayahuita with ‘esophagus.’ Miskito *won krukmaya* consists of *krukmaya* ‘goiter’ and the classifier *won*, and Wayampi colexifies the relevant

meanings directly. Kyaka colexifies ‘Adam’s apple’ with ‘trachea,’ and Wintu *holhol* is reduplicated from *hol* ‘pipe, tube, trachea’ (also meaning ‘light, bright, shine’).

Other associations include: Dongolese Nubian *gōz* also means ‘dune,’ though in this meaning it is borrowed from Sudanian Arabic. Dadibi *dogoni* also denotes the ‘sound of voice,’ and the Kwoma term *noku pipoy* appears to be analyzable as ‘sago.palm strike.down.’ Abzakh Adyghe *qʷəmbəqʷay*, colexifying ‘glottis,’ contains either *qʷə* ‘skeleton’ or *(-)qʷə* ‘pointed object,’ the epenthetic element *-m-*, and *b(e)*, which occurs in a number of complex terms referring to the human body. Badaga *gəttigallu* is analyzable as */gətti-kallu/* ‘cartilage-stone,’ and Cheyenne *-nêhpoʼhooʼôtse* as */-nêhpoʼohe-hestôtse/* ‘locked-thing’ (the literal translation given in the source is ‘valve/thing that closes’). The Chickasaw term *inonkopoolo* is perhaps analyzable as */inonkaʼ oppolo-ʼ/* ‘voice be.ruined-NMLZ.’ Piro *swayipla* appears to contain *swa* ‘orifice,’ Toba colexifies ‘Adam’s apple’ with ‘gill,’ and the Yanomámi term *krukupi* is also used to refer to the ‘hyoid bone of the Araguato monkey.’ Fijian *i tagi-tagi* is analyzable as ‘DERIV make.sound-RED,’ Hani *kaoqciivq* might be related to *ciivq*, meaning ‘to choke, pinch’ and ‘genealogy,’ alongside acting as a classifier for ‘joints,’ for instance on the fingers or on bamboo. Samoan colexifies ‘Adam’s apple’ with ‘eyes of a snake or eel’ and ‘uvula.’

#### 99. The Ankle

Representation: 86%

Motivated: 46.1%

Thereof Analyzable: 41.7%

Thereof Colexifying: 6.4%

Thereof by Contiguity: 13.5%

Thereof by Similarity: 29.2%

Recurrent associated meanings: foot/leg, joint/wrist, eye, neck/throat, knot, bone, egg, heel, (Achille’s) tendon, stone/pebble, seed

Motivated terms for ‘ankle’ are overwhelmingly morphologically complex, with ‘foot’ and/or ‘leg’ being the meaning of one of the elements. Mostly, the semantics of the second element is such that it denotes a small roundish object, thus constituting a metaphorical transfer. However, this is not always so clear, and in fact the most frequent subtypes attested in the sampled languages do not adhere to this pattern.

In thirteen sampled languages, Swahili, One, Toaripi, Badaga, Lake Miwok, Kaingang, Miskito, Wayampi, Hawaiian, Kapingamarangi, Mandarin, White Hmong, and Bislama, ‘joint’ or ‘wrist’ is the meaning of the second constituent (the Hawaiian term colexifies ‘heel,’ as does a simplex Badaga term). For one, One has *teu tampla* ‘foot/leg joint;’ Welsh, Cahuilla, Haida, Wintu, Rama, and Tehuelche colexify ‘wrist’ or ‘joint’ with ‘ankle,’ Rama also colexifies ‘knuckle’ specifically. The next most frequent subtype is that where the meaning of the second constituent is ‘neck’ and/or ‘throat,’ as in Sahu *rou ma camala* ‘foot/leg POSS neck,’ clear cases of which are found in nine sampled languages, alongside Sahu in Efik, Yoruba, Sko, Abzakh Adyghe (where ‘neck’ is colexified with ‘top’), Japanese, Itzaj, White Hmong, Vietnamese, and possibly Bakueri, though this is unclear (the Sko term stands out in that here not ‘neck,’ but ‘neck bones’ specifically is the meaning of the constituent, perhaps providing a clue to the underlying perceived similarity responsible

for this strategy in general). There are also associations with ‘bone’ in other languages: Santiago Mexquititlan Otomí has *bots'undo'yo wa*, containing *ndo'yo* ‘bone’ and *wa* ‘foot,’ Rotokas and Cubeo have terms with a similar structure and Cubeo also others, in which the motivation is constituted by classifiers; Copainalá Zoque has a semianalyzable term where the identifiable constituent is bone.’

A more clearly metaphorically motivated pattern of transfer from one (“salient”) body-part to another (less “salient”) one is constituted by complex terms where the meaning of the second constituent is ‘eye,’ as in Koyraboro Senni *čee-moñe* ‘foot-eye.’ Such terms are also found in Hausa, Kanuri, Anggor, Kwoma, Kyaka, Kosarek Yale, Pipil, and Hani. Moreover, Piro *yhale-xi* is analyzable as ‘eye-fruit/seed/DIM’ (the term colexifies, in addition, ‘leaf bud’ and ‘belt buckle’), and Sedang directly colexifies ‘ankle’ with ‘eye’ inter alia. Again, one term stands out, namely that in Kosarek Yale, where an additional constituent meaning ‘fruit,’ ‘seed,’ and ‘egg’ inter alia (see § 6.4.3.13.3. for this pattern of colexification) is present. A semianalyzable term with ‘fruit,’ colexifying ‘kidney,’ ‘side of ankle’ and ‘earlobe,’ is also found in Lenakel, and an association between ‘ankle’ and ‘egg’ is also revealed by the Chayahuita term *ca'yo-pira* ‘egg-pile’ and is also present in Bwe Karen, which has *-khañi di*, containing *-kha* ‘leg, foot’ and *di* ‘egg, testicle.’ Common are also complex terms with ‘foot’ and/or ‘leg’ being the meaning of one constituent, and ‘knot’ that of the other, as in Badaga *gañnu kalu* ‘knot/joints leg.’ Such terms (with or without colexification with ‘joint’) are also attested for Pawnee, Tsafiki, and Tetun, direct colexification is attested in Tehuelche, and a semianalyzable term with one constituent meaning ‘knot’ or ‘joint’ is found in Rotuman. Further, Dongolese Nubian has *óssiñkugúndi* ~ *óssiñkumúndi* /*óssi-ñ-kugúnd(i)*/ ‘foot-GEN-elbow,’ Aymara has *kayu moqo* ‘foot pebble,’ and Arabella colexifies ‘ankle’ and ‘stone, rock.’ Further complex terms with one of the constituents being ‘foot’ or ‘leg,’ many of them reconcilable with the general tendency of metaphorical transfer of small roundish objects, are Yoruba *kókosè*, analyzable either as /*kókò-è-sè*/ ‘edible.tuber-leg’ or /*kókó-è-sè*/ ‘knot.in.tree/knob/excrescence-leg,’ Sko *làngòde*, presumably analyzable as /*làng-óe*/ ‘foot-yam,’ Abzakh Adyghe *leč'en* /*λ(e)-č'en*/ ‘foot-ossicle/lump’ (a semianalyzable term containing an element which can refer to a ‘lump’ inter alia is found in Rotuman), Biloxi *sponi'* ~ *i'sponi'* ~ *iñksponi'* ~ *sponitu'* ~ *i'sponitu'* ~ *iñksponitu'* /*asi-po-ni'*/ ‘feet-wrap.up.in.bundle-CAUS,’ Chickasaw *iyymosak* /*iyyi'* imosak/ ‘foot DAT-hickory.nut’ (the analysis is marked as questionable in the source), Highland Chontal *galgoxac gahmis* /*galgoxac lahmis*/ ‘shelled.corn leg,’ Kashaya *q<sup>h</sup>ama p<sup>h</sup>iko* ‘foot ball,’ Imbabura Quechua *chaki muku* ‘foot nude,’ Wayampi *piyũ?ã* /*pi-yu-?ã*/ ‘foot-spine-maintain.upright.position’ (this term colexifies ‘claw of bird of prey’), Bwe Karen *-kha-dé* ‘-foot-narrowest.part,’ Samoan *tapu-vae* ‘forbidden-leg/foot’ (colexifying ‘Pig’s trotters’), Takia *ñe-n ñdu-n* ‘leg-3SG nose-3SG,’ Tetun *ain-liras* ‘foot/leg-wing/fin,’ and Yay *ho'ðay' tin'* ‘adam’s.apple foot.’ Semianalyzable terms of this kind are found in Mbum, Kaluli, Kemtuik, Ngaanyatjarra, Rotokas, Sentani, Yir Yoront, Badaga, Carrier, Kiliwa, Kiowa, Oneida, San Lucas Quiaviní Zapotec, Guaraní, and Hani. Aguaruna colexifies ‘ankle’ and ‘tendon,’ and Yir Yoront colexifies ‘Achille’s tendon’ specifically (Guaraní, in addition, has a semianalyzable term for the ‘ankle’ where the meaning of the identifiable constituent is ‘Achille’s tendon’).

Other associations include: Khoekhoe *!goros* also denotes the ‘fetlock joint of animals,’ and Rendille, presumably due to homonymy, colexifies the “self, human ego.” Ngaanyatjarra *warungantjil(pa)* also denotes a “solid log that continues burning a long time” and the ‘Australian Hobby’ (a species of falcon). Rotokas *arioisi* also means ‘kapiok seed.’ Khalkha *šaya(n) ~ siya(n) ~ siyai* may be related to *šaya- ~ siya-* ‘to hit with fist, drive a nail or peg in.’ Welsh *swrn* also means ‘fetlock’ and ‘a good number.’ Ineseño Chumash *šow* colexifies ‘ankle’ with ‘tobacco,’ while Lesser Antillean Creole French *chivi* also means ‘plug, peg’ and ‘pin.’ Central Yup’ik *cingilleq* is derived from *cingiq* ‘shoelace’ by means of the postbase (see § 4.4.2) *-lleq* ‘former.’ Bora *lleébou* might be related to a verb meaning ‘to listen, obey,’ Cayapa *nemiipijpuca* contains *puca* ‘small round thing,’ while Macaguán *peatá* appears to be derived from *atá* ‘hard’ by prefixation of a possessive marker. The Jarawara term *rabi* also means ‘pencil’ due to Portuguese influence, while Ancash Quechua *utsu putu* is analyzable as ‘garlic container’ and indeed also denotes a vessel in which garlic is stored and which is similar in appearance to the ankle. Toba *nqonacamo* appears to contain *qona* ‘toe, claw,’ and Malagasy *kitro kèly* is analyzable as ‘hoof small.’

#### 100. The Beard

Representation: 92%

Motivated: 51.8%

Thereof Analyzable: 33.5%

Thereof Colexifying: 19.9%

Thereof by Contiguity: 27.5%

Thereof by Similarity: 15.9%

Recurrent associated meanings: hair, chin, mouth, whiskers, fur, feather, jaw,  
lip, antenna, beard of grain, face, corn silk, barbel, wool

The most frequently associated meaning for ‘beard’ (without differentiating between ‘beard,’ ‘mustache,’ ‘goatee,’ etc.) is, unsurprisingly, ‘hair.’ Kosarek Yale, Tasmanian (Northeastern), Nivkh, Bororo, and Ancash Quechua colexify ‘hair’ and ‘beard’ directly; otherwise, the association is realized by complex terms, the other element having a variety of meanings in the languages of the world. Most frequently, in Mali, Toaripi (for ‘mustache’ specifically), Sora, Cheyenne, Kiliwa, Nuuchahnulth (with the element referring to ‘mouth’ being a lexical affix), Pawnee (where ‘be hairy’ is colexified with ‘be furry, fuzzy’), Pipil, Wintu, Copainalá Zoque, Lenakel, Malagasy, and Tetun, the second constituent makes reference to the ‘mouth,’ as in Tetun *nunun-rahun* ‘mouth/lip-body.hair,’ which also denotes a ‘goatee’ specifically (the constituent means ‘hairy’ rather than ‘hair’ in the case of Cheyenne; this situation is recoverable etymologically in Cahuilla, compare also Hani *meiqmoq* with *meiqbaoq* ‘mouth, snout’ and *aqmoq* ‘horse, pubic hair’). There are, however, also complex terms of the lexical type with ‘mouth’ where the second element has a meaning other than ‘hair.’ Khoekhoe has *lhō-ams* ‘wrinkle-mouth,’ Ngambay *mbáy-tà* ‘lord-mouth/lips,’ Abzakh Adyghe *žac’že* /že-č’že/ ‘mouth-end,’ Cashinahua *kex-ni* ‘mouth/lip/edge-forest,’ Hupda *nə-cúg* ‘mouth-hummingbird,’ and Takia *awa-n dabi-n* ‘mouth-3sg root-3sg’ (it is possible that clearly metaphor-based terms as in the latter two languages are due to dense beards being an item of “acculturation” so to speak - growth of bodily hair is less pronounced for instance among the indigenous population of the Amer-

icas when compared with Europeans). In Buin and Yir Yoront, there are semianalyzable terms for 'beard' featuring a constituent with the meaning 'mouth.'

Also very common are terms of the analyzable type with the constituents meaning 'hair' and 'chin,' as in Berik *olele safa* 'chin hair/fur/feather.' Such terms are also attested in Efik, Dadibi, Kwoma, Meyah, Toaripi, Lakhota, Rama, Toba, Wayampi, Great Andamanese, and Bwe Karen. Other complex terms on the basis of 'chin' are Katcha *musá mo buruŋe* ~ *musá mo böröŋe* 'wool GEN chin' (similarly, the association with 'wool' is present due to colexification with 'hair' in the Toba term mentioned above), Waris *keu-ta* 'chin-small.object,' and Bora perhaps *újca-he* 'chin/beard-CL.oblong,' but segmentation is unsure. 'Chin' and 'beard' are directly colexified in Dongolese Nubian (also with 'whiskers,' a pattern of colexification also attested in Baruya, Nunggubuyu, Rotokas, Western Tasmanian, Carrier, Chickasaw, Ineseño Chumash, Haida, Kiowa, Lake Miwok, Pawnee, Xicotepec de Juárez Totonac, Wappo, Wintu, Central Yup'ik, San Lucas Quiavini Zapotec, Bislama, where it is archaic, and Hani, where 'chin whiskers' more specifically is colexified), Gurindji, Ngaanyatjarra, Northeastern, Middle Eastern, and Western Tasmanian, Badaga, Embera, Fijian, Rotuman, and Tetun; a semianalyzable term where the identifiable constituent is 'chin' is featured in Chukchi and Middle Eastern Tasmanian (the association is common in Indo-European, Buck 1949: 205). Kwoma and Rama colexify 'chin' with 'jaw,' and therefore their terms for 'beard' also feature an association with 'jaw' more specifically. Complex terms with 'jaw' alone and 'hair,' such as Yaqui *chao boa-m* 'jaw body.hair-PL,' are also found in Yoruba, Sko, and Chickasaw. Miskito colexifies 'beard' with 'jaw' (perhaps a compound with 'jaw' is present in Ancient Greek, Buck 1949: 205). Baruya has *maryaata* /*marya-mijata*/ 'lip-hair,' and terms with such structure are also found in Ngambay, Abipón, Piro (where the constituent making reference to the lips means 'rounded lips' more precisely), and Tetun, as already noted above, while Cashinahua has, as also already noted above, *kex-ni* 'mouth/lip/edge-forest.' Jarawara has *noki baki kone/noko bako kone* 'face hair' for 'beard on side of face,' and terms with such structure (without a semantic restriction mentioned like that in Jarawara) occur also in Nuuchahnulth (with the 'face' component expressed by a lexical affix) and Wintu (where 'face' and 'eye' are colexified). Other complex terms where one of the constituents is 'hair' are Kiowa *sęin-p'ā'-ga* 'mucus.of.nose-body.hair-NOUN.POSTFIX' and Hawaiian *hulu weuweu* 'body.hair grass/bushy' for "[d]owny feathers or beard, fuzz." A semianalyzable term is found in Japanese.

Now, the above discussion does not take into account the lexical differentiation made in many languages between 'hair of the head,' 'body hair,' and potentially even more fine-grained distinctions. Since 'body hair' is frequently colexified with 'fur of animal' and/or 'feathers,' a number of languages betray associations also with these meanings. With 'fur,' this is the case by colexification in Kosarek Yale and Ancash Quechua and by analyzable terms in Berik, Kwoma, Meyah, Toaripi, Chickasaw, Kiliwa, Lakhota, Nuuchahnulth, Pawnee, Wintu, Toba, Bwe Karen, Lenakel, and Tetun. In two sampled languages, Kolyma Yukaghir and Kashaya, it is specifically 'fur' rather than 'hair' which is the constituent of complex terms. Note also Tehuelche *ʔašč'ex* ~ *ač'ex* ~ *ʔačx* ~ *ʔač'ex* ~ *ʔašč'ex* ~ *ač'ex* ~ *ač'ex* 'in-hair/fur,' which colexifies 'beard' with 'eyebrow' and 'eyelash.' 'Beard' is associated by analyzable terms with 'feather' in Efik, Berik, Kwoma, Meyah, Toaripi,

Chickasaw, Toba, Bwe Karen, Hawaiian, Lenakel, and by colexification in Kosarek Yale and Bororo (sometimes also with further meanings, cf. section 21).

Furthermore, Xicotepec de Juarez Totonac, Arabela, Cavineña, and Piro colexify ‘beard’ with ‘antenna of insect,’ Basque, Khalkha, and Nez Perce with ‘beard of grain,’ Basque and Hawaiian with ‘barbel,’ and Kiowa and Piro with ‘corn silk’ specifically.

Other associations include: Hausa *gemu* also denotes the concepts ‘corbel’ and ‘superior person,’ while Koyraboro Senni *kaabe* ~ *kaaba* is also the name of a spice and a tree species. Gurindji *jika* is also the term for the ‘brush-tailed rat kangaroo’ and ornaments made from its tail. Kwoma colexifies beard with “shell and feather decoration glued on the jaw of certain ceremonial sculptures carved in the form of heads,” and Muna has *samba* ~ *kasamba* for “beard (on chin and cheeks),” which is borrowed from Indonesian; indigenously, *samba* also denotes a kind of girdle inter alia. Khalkha *ziber* also means ‘fins of fish’ (the meaning ‘beard’ is register-specific), Nivkh colexifies ‘beard’ with ‘vegetation,’ ‘feeler,’ and ‘tentacle’ (the term is very short, so homonymy is a possibility). Cahuilla *támaš* is also used to refer to a bearded person. Santiago Mexquititlan Otomí colexifies ‘beard’ and ‘flour,’ while Guaraní *tendyva* contains *tendy* ‘saliva.’ Ancash Quechua *sapra* is also metonymically capable of referring to a ‘Spaniard.’ Fijian *kumi* also denotes ‘Tongan cloth,’ Kapingamarangi *tuu i lala* seems to be analyzable as ‘cut at down,’ while Rotuman *kumkumu* also denotes a species of crab. Bislama *mustas* (< Engl. *mustache*) is also the name for the ‘goatfish.’

#### 101. *The Bladder*

Representation: 59%

Motivated: 67.2%

Thereof Analyzable: 47.3%

Thereof Colexifying: 19.9%

Thereof by Contiguity: 31.6%

Thereof by Similarity: 22.4%

Recurrent associated meanings: urine/urinate, container, bag, balloon, place,

blister, gall bladder, bubble, belly/stomach, sack, house, goiter, eye, inflate

Unsurprisingly, terms for the ‘bladder’ frequently contain an element meaning either ‘urine’ or ‘to urinate.’ ‘Urine’ / ‘urinate’ and ‘bladder’ are colexified directly in Anggor, Ngaanyatjarra, Tuscarora, Aguaruna, Huambisa, Yanomámi, and Bislama. Derived terms, such as Ineseño Chumash *šoxšol*, derived from *oxšol* ‘to urinate,’ are also found in Bakueri, Hausa, Upper Chehalis, Abipón, Bora, Chayahuita, Cubeo, and Toba. In terms of the lexical type, ‘container’ is most frequently the meaning of the second constituent, as in Ket *dysol* /*d̥s-óðl*/ ‘urine-container,’ and also in Chickasaw, Kashaya, Kiliwa, Nuuchahnulth, Bororo, Carib, Kaingang, and Ancash Quechua. Similarly, the Toba term is derived by a classifier for receptacles. A specific container, namely ‘sack,’ is also attested as the meaning of the second constituent, as in Mbum *fǎ-jáu* ‘sack-urine.’ Such a term is also found in Ket, and ‘sack’ and ‘bladder’ are colexified directly in Nez Perce inter alia. Very similarly, Yoruba has *àpo-ìtò* ‘bag-urine,’ and so do Carrier and Kiowa, while the association is realized by colexification in Piro and Hawaiian (furthermore, Tuscarora has an optional complex term of this kind on the basis of its term colexifying ‘bladder’ and ‘urine’).

Semianalyzable terms where the identifiable constituent means ‘sack’ or ‘bag’ are found in Highland Chontal and Samoan. Terms with ‘house’ (which is a kind of container of sorts as well), such as Embera *súa-dé* ‘urine-house,’ are also attested in Mbum and Miskito (in the latter language, there are also alternative terms where ‘house’ is replaced by ‘womb’ and ‘urine’ by ‘liquid’ as a constituent). More generally, ‘place’ is the meaning of the second constituent in Arabela *shaaca quiquio* ‘urine place,’ as well as in Lake Miwok, Miskito, and Great Andamanese. Rama *ngústi úp* is analyzable as /ngústi up/ ‘pissing eye,’ while Kyaka has *puu renga* ‘urine eye’ (both *puu* and *renga* have other meanings). Other complex terms where ‘urine’ is a constituent are Buli *sinsam-liuuk* ‘urine-passage/hole/path,’ Efik *u-tök ikim* ‘NMLZ-to.void urine,’ Katcha *nimo misəɔ* /nimo m-kisəɔ/ ‘thing ??-urine,’ Cheyenne *xaenoo’ôtse*, analyzable as /xae-noo’ohtsé/ ‘urinate-depart,’ Oneida *-nhahálakhwa?*, containing the roots *-nhah-* ‘urine’ and *-l-* ‘be in or on,’ Pipil (Cuisnahuat dialect) *-a:xix-tekun* ‘-urine-bottle.gourd,’ Yuki *as’-pis’* ‘urine-from,’ San Lucas Quiavini Zapotec *gui’dy x:qui’ix* /gui’ihdy x:qui’ix/ ‘skin urine,’ Tehuelche *t’ep’en?aşk’en*, containing *t’ep ~ t’e:p ~ t’ep’ ~ t’epe ~ ?et’ep* ‘urine, urinate’ and *-aşk’en* ‘interior,’ and Hawaiian *pu’u-mimi* ‘protuberance-urine.’ Semianalyzable terms are found in Bakueri, Kaluli, Chukchi, Biloxi, and Guaraní. Wintu *p<sup>h</sup>u’lemes* and *p<sup>h</sup>u’les* contain *p<sup>h</sup>uł* ‘to blow,’ the former colexifying ‘bologna, meat in casings, sausages’ (note also Fijian *i uvu-uvu* ‘DERIV blow.with.mouth-RED’).

Basque, Embera, and Macaguán colexify ‘bladder’ with ‘blister,’ Ngambay and Hawaiian with ‘goiter’ (Ngambay also with ‘evacuate’ and ‘gall bladder,’ it shares the latter association with Ngaanyatjarra and Yir Yoront, and Hawaiian also colexifies ‘womb’ among other meanings), while Pawnee and Cavineña have terms derived from verbs meaning ‘to inflate’ (*paaruris*, analyzable as /waarur-his/ ‘inflate-PERF’ and *cujacuja*, reduplicated from *cuja* ‘to inflate’ respectively; both terms colexify, as do Lesser Antillean Creole French and Samoan, ‘balloon’). ‘Bladder’ and ‘stomach’ are colexified in Nuuchahnulth and Hawaiian, and similarly, Lenakel has *nouanetpinami* /noua-netp-nami-/ ‘fruit-belly-urine-,’ which colexifies also ‘ball.’ Finally, Itzaj, Cayapa, and Hani colexify ‘bladder’ with ‘bubble’ (Cayapa colexifies ‘animal bladder’ more specifically, also with ‘foam’).

Other associations include: Buin *kou* is a general term for the internal organs such as ‘guts,’ including the bladder. Kwoma *mokugwey* appears to contain *moku* ‘semen,’ and Chukchi *pəyəlqewəc?en* contains *pəyəl-* ‘float.’ Greek *kýstis* also means ‘cyst,’ and Khalkha *dabusang ~ dabasay ~ dabisay* is also used to refer to the ‘lower part of the abdomen’ and ‘the pubic region.’ San Mateo del Mar Huave *mipeparan* seems to contain the verb *apep* ‘to catch.’ Itzaj *b’ejiiqaj* also means ‘innertube,’ while Nez Perce colexifies ‘bladder’ with ‘shell’ and ‘coffin.’ Xicotepec de Juárez Totonac *a’nlhā tāquī* *xcān* appears to be analyzable as ‘where get.up water,’ and Central Yup’ik *nakacuk* might contain the postbase (cf. § 4.4.2.) *-cuk* ‘unpleasing, bad one.’ The Aymara term *llaq’allachi ~ yaq’allachi* may contain *yaq’a* ‘excrement,’ while Guaraní *tyryru* also means ‘urinal.’ Hawaiian *ko’ana* also is used with the meaning ‘dregs, sediment.’

102. *The Blood*

Representation: 98%

Motivated: 15.4%

Thereof Analyzable: 2.4%

Thereof Colexifying: 13.0%

Thereof by Contiguity: 7.4%

Thereof by Similarity: 1.5%

Recurrent associated meanings: red, resin, race/kinship, menses/menstruate, fetus, body liquid, liquid

Terms for ‘blood’ are not frequently motivated. The most frequent association (ignoring ambiguous terms with respect to word class which can hence also mean ‘to bleed’), occurring in nine sampled languages, is that with the color ‘red,’ by colexification in One, Toaripi (“blood red” more specifically), Wintu (colexifying also “arse”), Aymara, Rama, and Hawaiian, and by complex terms in Abipón and Bora (*n-aoig-Ra* ‘POSS.INDEF/3SG-red/yellow-CAUS’ and *tújpa-cyo* /*tújpá-cyo*/ ‘red-CL.pointed’ respectively). Yanomámi *iyěiyě* ~ *iyěiyě* ~ *iyěiyě* is reduplicated from *iyě*, which means ‘bloody,’ ‘blood-colored,’ and ‘red.’ The association with ‘red’ is also found in Sanskrit (Buck 1949: 206). A semianalyzable term is furthermore found in Yana, and the association is etymologically recoverable also in Dongolese Nubian. Colexification with ‘resin, sap’ is found in Kwoma, Cubeo, Maxakalí, Fijian, and Hawaiian (note also *Sko hi* ‘blood’ and *hi* ‘sap’ and the colexification of ‘blood’ with “red sap of certain trees” specifically in Jarawara). ‘Blood’ and ‘body liquid’ generally are colexified in Kyaka and Hawaiian (in Kyaka, the term also denotes ‘lymph’ and ‘serum,’ and, presumably by homonymy also means “song, singsang, chant”), and, still more generally, ‘blood’ and ‘liquid’ are in Bororo (among other meanings), Maxakalí, and Hawaiian. In Wayampi, *tieʔe* contains *tii* ‘juice, liquid,’ and the other element is said to etymologically have the meaning ‘true, genuine.’ Another Wayampi term, *tuwi*, also means ‘dry.’ Furthermore, ‘blood’ is colexified with ‘menses’ or ‘to menstruate’ in Dongolese Nubian, Lake Miwok, and Samoan (where the relevant term is reduplicated from *ʔele* ‘red soil’ and also denotes the ‘earth’ and the ‘dirt’), with ‘race’ or ‘kinship’ in Basque, Lesser Antillean Creole French (inter alia), Itzaj, Ancash Quechua, and Rotuman, and with ‘fetus’ in Efik and Tetun.

Other unique associations in the languages of the sample include: Efik *iyip* is also used with the meaning ‘murder, bloodshed,’ and Hausa *jini* is also the name of a children’s game. Kwoma is unique in using the same term, *pi*, for ‘blood’ and ‘milk’ inter alia, and Muna *rea* also means to “have a communal garden.” Waris *tóvól* ~ *nihtóv* also means ‘skin,’ and Kosarek Yale *eneng* also ‘pus.’ Abzakh Adyghe *λə* also means ‘be lying, suspended, situated’ inter alia, and Haida *ray* is also used with the meaning ‘kidney of salmonid.’ For *utkwàreh* ~ *útkwareh*, a Tuscarora term for ‘blood,’ the lexicographer remarks that the underlying root *-tkwar-* ~ *-tkwār-* might be an old compound of *-tkw-* ‘stomach’ and *-r-* ‘to be in.’ Central Yup’ik *qayyuq* also means ‘soup, broth,’ and Copainalá Zoque *napin* probably contains *na* ‘water.’ Toba *ltago’q* also means ‘one’s own son.’ Bwe Karen colexifies ‘to steam’ inter alia and by another term also ‘nest,’ and Hawaiian *koko* also means ‘rainbow-hued’ inter alia.



103. *The Bone*

Representation: 99%

Motivated: 21.7%

Thereof Analyzable: 0.7%

Thereof Colexifying: 20.3%

Thereof by Contiguity: 6.1%

Thereof by Similarity: 11.0%

Recurrent associated meanings: skeleton, stalk/fibre, seed, strength, awl/needle,  
hard/hardness, leg, tree, horn, fishbone, shell, dry, corpse

Like 'blood,' 'bone' is a meaning for which few lexical associations are found cross-linguistically. The most frequent one, by configurational contiguity, is that with 'skeleton,' occurring by colexification in Khalkha, Tuscarora, Bororo, Cashinahua, Embera, Guaraní, Huambisa, Tehuelche, and Rotuman. Further, Abzakh Adyghe has *q<sup>o</sup>ǎ-mǎ-shē* 'skeleton-RELAT/EPEN-upper.part/??.' Khalkha inter alia colexifies 'bone' with 'corpse,' and Hawaiian with 'carcass,' as of chicken. In Wappo and Bororo, relevant terms also mean 'hard' or 'hardness' (note also Rotokas *kerua*, presumably containing *keru* 'harden like bone' and a classifier for narrow objects and Sentani *bo* 'bone' and *bobo* 'hard'). Cheyenne and Lengua colexify 'bone' with 'leg' (note also the cognacy of German *bein* 'leg' and English *bone* as well as a further association of this kind in Greek noted by Buck 1949: 207; perhaps this pattern is motivated by the fact that the thigh bone is the largest bone in the human body?). Wappo and Yuki, presumably by functional contiguity, use the same word for 'bone' and 'awl' or 'needle.'

There are also a number of metaphor-driven associations, although these are not very frequent as well. Baruya, Ngaanyatjarra, Toaripi, Sahu, Sora, Arabela, Lengua, Maxakalí, and Yanomámi colexify 'bone' with 'stalk' or 'fibre' (Yanomámi also with 'cotton' and 'animal fur'). This pattern is also found in Indo-European as evidenced by cognates in Baltic, Greek, and Latin (Buck 1949: 207). Maxakalí and Tsafiki colexify 'bone' with 'tree' (as well as 'stick' or 'pole' in Tsafiki), Yir Yoront, Basque, Khalkha, and Wintu (among other meanings in the latter two languages) with 'seed' (this pattern is also attested in Indo-European, Buck 1949: 207), Highland Chontal and Itzaj with 'horn' (the latter language also with 'employment'), and Buli and Basque with 'fishbone.' San Lucas Quiaviní Zapotec *zu'aht* also denotes the 'shell,' e.g. of a turtle (alongside a rattle used in a particular dance), and Hawaiian *iwi* also colexifies 'bone' with 'shell,' among other meanings. In Ngambay, Baruya, Kwoma, and Waris, relevant terms also mean 'strong' or 'strength,' see Aikhenvald (2008: 579) for brief discussion of this association in Manambu suggesting areal convergence. In Anggor, *hamindi* is also glossed as 'very,' which suggests that it is used as an intensifier. Finally Guaraní *kâ* also means 'dry' (and 'pit of fruit'), and Burarra (-)jorla, meaning 'bone' in the Gun-nartpa dialect, generally also means 'dry' (as well as 'sun-baked' and 'no water').

Other associations include: Efik *ök'pö* also denotes a kind of yam, the fruit of a particular tree, a creeping vine, and 'bird lime.' The Burarra term *-mama* also denotes the 'frame of a structure' and a 'coin' as opposed to paper money. Dadibi *dili* also means 'singsing.' Gurindji *kuyuwarn* contains *kuyu* 'meat,' whereas the Kwoma term *hapa* also means 'bone of the upper arm' specifically and then by extension

also ‘upper arm.’ The Ngaanyatjarra term *tarrka* also means ‘bare area’ as well as ‘skinny’ in an adjectival sense. Greek *kókkalo* also may refer to ‘ivory on the piano key’ and ‘shoe-horn,’ and Ket *aʔt* also means ‘soon.’ As a verb, Kiowa *t’ēi’m* means ‘to break.’ Lake Miwok *kúlum* also means ‘cemetery,’ and Macaguán *-chit* also ‘shinbone.’ Hani *saqyyuq* and *saqyoq* may be related to *saq*, meaning ‘muscle, meat’ inter alia, Bwe Karen *khwi* is also the name of a sand lizard, the Kapingamarangi term *iwi* also means ‘upper ridge pole’ as well as ‘crease’ and ‘sharp edge,’ and the cognate Samoan *iwi* also means ‘limb.’ Rotuman colexifies ‘bone’ with “projecting spike ... on the fins of certain fish” inter alia, Takia *tatu* also means ‘spine,’ while Tetun *ruin* also means ‘rough,’ and Sedang *koxiang* may or may not contain *koxi* ‘egg.’ Yay *dok*<sup>2</sup> colexifies ‘flower’ and *ku*<sup>5</sup> also means “pair, even (of numbers).”

#### 104. *The Brain*

Representation: 90%

Motivated: 33.8%

Thereof Analyzable: 20.2%

Thereof Colexifying: 13.7%

Thereof by Contiguity: 9.0%

Thereof by Similarity: 22.2%

Recurrent associated meanings: head, marrow, mind, pus, semen, pith, faeces

Frequently, motivated terms for the ‘brain’ are morphologically complex and contain a constituent meaning ‘head’ (see Buck 1949: 215 for similar evidence from Indo-European). As for the semantics of the second constituent, ‘marrow’ is well attested. For instance, Kaingang has *krĩ kujo* ‘head marrow,’ as do Koyraboro Senni, Berik, and Kiliwa. ‘Brain’ and ‘marrow’ are, however, also frequently directly colexified. This is the case in Yoruba, Buraarra, Kyaka, Yir Yoront, Chickasaw, Central Yup’ik, Arabela, Bororo, Ancash Quechua, Toba, Yanomámi (where *heoshipě* ~ *heyōshipě* appears to contain *heōshi* ‘hot’ and the quantal classifier *pě*, for which see § 4.4.1.1.), Hawaiian (alongside other meanings), Lenakel, White Hmong, and Yay, while Bwe Karen *í*, colexifying ‘to give,’ is said to appear to be an alternant of *i* “soft centre (of a plant or tree), marrow (of a bone), etc.” Furthermore, Great Andamanese has a derived term from a root colexifying ‘marrow’ and ‘pus’ (the association with ‘marrow’ is not at all unknown, see e.g. Buck 1949: 215 and Jóhannesson 1949: 88 for Indo-European, Matisoff 1978: 233 for Southeast Asian, and Wilkins 1996 for Australian evidence; note in this context also the quote from Aristotle mentioned by Buck 1949: 215: “for many think the brain is really marrow”). Great Andamanese, in fact, colexifies ‘pus’ and ‘marrow,’ and an association with ‘pus’ is also found for ‘brain’ by colexification in Ket and Lake Miwok and by the analyzable terms *wi-dò* ‘pus-head’ in Ngambay and *lal mabiara* ‘head pus’ in Miskito (note also Kiliwa *?iy=xtla=?pii* ‘head=?=pus,’ with ‘pus’ a mistake for ‘pus?’). Basque has *burmuin*, presumably analyzable as /buru-muin/ ‘head-pith/bud/shoot/sprout’ (note that there is a complex term on the basis of *muin* meaning ‘marrow’), and ‘brain’ and “pithy core of any tree” are colexified in Nuuchahnulth (alongside ‘heart’ and ‘spinal cord’). Other analyzable terms with ‘head’ being one constituent are Buli *zupuuk* /zuk-puuk/ ‘head-foam,’ One *sila silla* ‘head leaf,’ Abzakh Adyghe *šhekʷəpse*, containing *šhe* ‘head,’ *kʷe* ‘space, middle,’ and either *-ps(e)* ‘thread, string’ or *pse* ‘soul, life

principle' (there is also a variant with a different third element with unknown semantics), Kolyma Yukaghir *jo:n-qodo* 'head-lying,' Santiago Mexquititlan Otomí *ñäxumo* /ñä-ximo/ 'head-calabash,' Pawnee *paks-kiraar-u* 'head-liquid-NOM,' Chayahuita *motoro* /moto'-ro/ 'head-CLASS.PILE,' Cubeo *jipo-jia* 'head-CLASS.RIVER.LIKE.OBJECT,' Jarawara *tati afone/tati afone* 'head soft.core,' Rama *king kás* 'head meat,' Tehuelche *č'eter* *ʔašk'en*, containing *č'eter* ~ *č'eʔter* ~ *četer* 'head' and *-ašk'en* 'interior,' Malagasy *atidòha*, analyzable as /àty-n-lòha/ 'liver/center-GEN-head,' and Rotuman *uat p̄ar-p̄aru* 'head RED-mix.or.knead.with.hands.' Further, in both Greek and Welsh, terms for 'brain' are (etymologically) connected to 'head,' amended by a prefix meaning 'in' (Buck 1949: 215). Semianalyzable terms where the identifiable constituent is 'head' are found in Kaluli, Carrier, Haida, the Cuisnahuat dialect of Pipil, Carib, and Cavineña.

Another class of terms for 'brain' bears associations to cognitive processes. Efik, Laz, Lesser Antillean Creole French, Nuuchahnulth, San Lucas Quiaviní Zapotec, and White Hmong colexify 'brain' with 'mind' or like meanings, Guaraní colexifies 'brain' with 'understanding, judgment' and 'talent' and Basque with 'intelligence,' while Kashaya *hoʔto* contains *-oʔto* 'to think' and Miskito *won sinska* contains *sins* 'intelligence.' A further language where the word for 'brain' appears to be related to cognitive abilities is Toba, but the precise structure remains unclear.

Further, Ket and Lake Miwok colexify 'brain' with 'semen.' An association restricted in the languages of the sample to the Barbacoan languages Cayapa and Tsafiki is that with 'faeces' (*mishpe* /mishu-pe/ 'head-excrement' and *fu-pe* 'hair-excrement' respectively).

Other associations include: Anggor *efu* is glossed also as "bee (sweat)," and Mali *genaing* also means 'phlegm, mucus,' while the Meyah term *ofóita* also means 'pulp.' Sahu *nyinyiala* appears to be derived by reduplication from *nyiala* 'canarium tree.' Badaga *mu:le* also means 'skull,' Ket *doʔŋ* also 'three,' and Khalkha *tariki(n)* ~ *taraki(n)* also 'occiput.' The Nez Perce term *yexyeqí-tes* is analyzable as 'tan-INSTR.' The underlying verb means more specifically "to put in a solution of brain as a step in tanning hide." Cashinahua *mapu* also means 'ashes,' 'soap,' the "head" of an axe, and is also used with reference to that part of a rudder where it is held. Guaraní *apy-tu'û* appears to be analyzable as 'extreme.point-softness,' and Ancash Quechua *tuqshu* ~ *toqshu* can also mean 'stupid.' Kapingamarangi *ngogo* also means 'egg' and 'zero,' Sedang *ngoa* might contain *ngo* 'mountain,' and Tetun *kakutak* is derived by the nominalizing circumfix *ka-...-k* from the verb *kuta* 'to smear.'

#### 105. The Breast

Representation: 98%

Motivated: 57.8%

Thereof Analyzable: 5.2%

Thereof Colexifying: 53%

Thereof by Contiguity: 32.5%

Thereof by Similarity: 5.6%

Recurrent associated meanings: milk, udder/teat, nipple/teat, suck, chest, heart, mammary gland, Burton's legless lizard, bay

The most common association for the ‘breast’ is that with ‘(mother’s) milk,’ occurring in as many as 49 sampled languages without clear areal tendencies, though the virtual absence of the pattern in Eurasia is notable. The relevant languages are Ngambay, Noni, Yoruba, Anggor, Baruya, Buin, Burarra, Gurindji, Kyaka, Lavukaleve, Ngaanyatjarra, Nunggubuyu, Toaripi (colexifying also ‘scrotum’ and denoting a particular breast-shaped shellfish), Sahu, Kosarek Yale, Basque, Kolyma Yukaghir, Carrier, Upper Chehalis, Cheyenne, Haida, Kiowa (colexifying also ‘to flow, melt’), Lake Miwok, Nez Perce, Nuuchahnulth, Oneida, Santiago Mexquititlan Otomí, Quileute, Xicotepec de Juárez Totonac, Wappo, Wintu, Yuki, Central Yup’ik (Yukon and Norton Sound dialects), San Lucas Quiaviní Zapotec, Aguaruna, Aymara, Cashinahua, Huambisa, Kaingang, Macaguán, Bislama, Hani, Hawaiian, Lenakel, White Hmong, Rotuman, Samoan, Takia, and Tetun; Hausa colexifies ‘breast’ with ‘sour milk’ specifically, in Swahili, the association is realized by noun class alternation, and in Kashaya *ši?do* is also the archaic term for ‘milk,’ while nowadays *molokko*, a loanword from Russian, is used.

The Haida term *tl'tn-ʔu* colexifying ‘breast’ and ‘milk’ is analyzable as ‘suck-INSTR,’ and so is the relevant Kolyma Yukaghir term. A derived term is also found in Rotokas, Sora (here, the term can dialectally also refer to the ‘mother’ herself), Upper Chehalis, and Tetun (and a semianalyzable one in Chukchi), while Muna, Basque, Itzaj, Lesser Antillean Creole French, Pawnee, Kapingamarangi, Samoan, and Bislama have terms directly colexifying nominal ‘milk’ with verbal ‘to suck,’ sometimes among other meanings. In Fijian *sucu* is ‘to be born, suck the breast, birth, milk,’ with milk also expressed by *wainisucu* (*wai* ‘water’), and ‘breast’ is *sucuna* (*-na* is a possessive marker).

The Nez Perce and Lavukaleve terms colexifying ‘milk’ and ‘breast’ at the same time also colexify ‘nipple’ and/or ‘teat’ (see Buck 1949: 248 for Indo-European evidence), and this is a pattern also found in Khoekhoe (here with ‘nipple of man’ specifically), Dongolese Nubian, Swahili, Yoruba, Northeastern Tasmanian, Badaga, Basque, Khalkha, Kolyma Yukaghir, Ineseño Chumash, Haida, San Mateo del Mar Huave, Itzaj, Lake Miwok, Nuuchahnulth, Pawnee, Pipil, Wintu, Aguaruna, Arabela, Aymara, Bora, Cavineña, Chayahuita, Embera (where the meanings are associated with different genders), Guaraní, Lengua, Miskito, Ancash Quechua, Toba, Yanomámi, Hawaiian, and Sedang, which colexifies also “to pound, to beat” (34 languages). Hausa and Kwoma colexify ‘mammary gland.’

Moreover, Cubeo *opebo* ‘breast’ might contain *ope* ‘nipple,’ but this cannot be determined with certainty on the basis of the consulted source. In many sampled languages, terms for ‘breast’ are also capable of referring to the homologous structure in animals, namely the ‘udder’ or ‘teat.’ This is found in Efik, Hausa, Khoekhoe, Dongolese Nubian, Yoruba, Kyaka, Toaripi, Badaga, Japanese (only in a term restricted to child language), Khalkha, Sora, Kildin Saami, Haida, San Mateo del Mar Huave, Itzaj, Lake Miwok, Nez Perce, Nuuchahnulth, Pawnee, Wintu, Yaqui, Aguaruna, Arabela, Aymara, Bora, Cavineña, Chayahuita, Embera (where the meanings are associated with different genders), Guaraní, Hupda, Lengua, Miskito, Pipil, Ancash and Imbabura Quechua, Toba, Yanomámi, Hawaiian, Malagasy, Mandarin, White Hmong, Rotuman, Takia, and Yay (44 languages). Hawaiian and Mandarin have optional complex terms on the basis of a term colexifying ‘milk’ and

'breast' that exhibit this pattern: *pu'u-waiū* 'protuberance-milk/breast' (which also colexifies 'wet, moist') and *ru3-fang2* 'milk/breast/drink-room' respectively.

Moreover, like English and other European languages, fifteen sampled languages, Sentani, Badaga, Greek, Khalkha, Welsh, Blackfoot, Cahuilla, Nuuchahnulth, Wappo, Bora, Embera (with gender differentiating between the meanings), Rama, Fijian, and Vietnamese, colexify 'breast' with 'chest' (Badaga also with 'arm'), and four sampled languages, Katcha, Highland Chontal, Kaingang, and Great Andamanese with 'heart' (Kolyma Yukaghir also with 'soul' and 'breath').

Further, in Burarra and Nunggubuyu, relevant terms for 'breast' and 'milk' are also used to refer to a species of lizard (Burton's legless lizard, *Lialis burtonis*), and in Basque and Greek, relevant terms can also refer to a 'bay.' Both associations are likely areal.

Other associations include: Efik *ēba'* is derived from *iba'* meaning 'two' among other things. Hausa *nono* also denotes "the fins below the head of a fish" as well as "a cluster of fruit," and *mama* is a "title and position among female girls" and also means "throwing pestle into the air each stroke when pounding." Dongolese Nubian *érti* has a homonym (?) meaning 'dirt, dirty,' Rendille *náhas* also means 'pump, waterpipe,' Yoruba *omú* also denotes the "instrument used in weaving to divide the woof," and Muna *titi* also means 'stalactite' inter alia. Sentani *nimə* also means 'ripe,' and Badaga *bo:si* also denotes the "act of fondling breasts" as well as "tumbler, glass" in the Kunde dialect, and *mai* can also refer to a "deep location." Abzakh Adyghe *žə* also means 'to throw, hurl,' Basque *golko* can also refer to the 'stomach, guts' and *bular* also means 'courage' inter alia. Japanese *chichi* also means 'father' (the meanings being distinct in writing). Khalkha *cegezi(n)* also may refer to the "memory as a faculty," and *elige*, figuratively, also means 'liver,' 'belly,' and "blood relatives," while Welsh *bron* is extended to 'breast of hill.' Lesser Antillean Creole French *sen* also means 'holy' and 'sane' inter alia, and Lake Miwok *téele* also 'to be slicing meat.' Nuuchahnulth *ʔinma* also denotes "Nob Point, where white powder seems to run out of a breast-shaped rock," and Pawnee colexifies 'breast' with 'body, corpse' inter alia. Wintu *ʔEm* can also mean "hold pectorally, carry something in the arms, embrace." Copainalá Zoque *cucpac* contains *pac* 'bone,' and *tzu'tzi* also denotes the "tooth of a stamp mill," while Bororo colexifies 'breast' with 'bud/shoot of plants' and Wayampi with 'stamen' of a flower. Macaguán *-apúchipar* ~ *-atbúchipar* contains *-chipár* 'finger,' Miskito *lama* also means 'present, benefit' and 'presence, proximity.' Piro *sta* colexifies 'front,' and there is also a verb *-sta-* 'to cut, tear, break.' Bislama *titi* and *susu* also mean "unweaned," as said of a child, and the former term also means 'to practice fellatio,' while Fijian *sere* also means 'to sit and sing' and 'to unloose, untie.' Hani *aqqul* also means 'sweet,' Hawaiian *ū* also 'moist, soaked' and 'to drip, drizzle, ooze' inter alia, and *waiū* also denotes a 'wet-nurse.' Kapingamarangi *uu* also means 'to absorb' inter alia. Finally, Rotuman *susu* also means 'to sew,' Samoan *mau*, a polite term for 'woman's breast,' also 'to keep, retain' and 'to live, dwell' inter alia, and Takia *su* also 'shoe' (with the additional sense < Engl. *shoe*).

106. *The Buttocks*

Representation: 84%

Motivated: 55.4%

Thereof Analyzable: 17.5%

Thereof Colexifying: 37.4%

Thereof by Contiguity: 22.4%

Thereof by Similarity: 17.9%

Recurrent associated meanings: bottom/base, back/behind, hip/haunch, anus, rump, end, tail, seat, cheek, faeces, stern of boat, thigh, hole, reason, root, sit

The most frequent association for the ‘buttocks’ is that with ‘bottom,’ occurring in 26 sampled languages (and also in English). There is a grammaticalization path said to be particularly common in African languages from ‘buttocks’ to ‘down’ (Heine and Kuteva 2002: 63); however, it is hard to impossible to tease apart these cases from those where relevant terms are glossed as ‘bottom’ as a slightly euphemistic alternative for ‘buttocks’ themselves. All cases are reported here in spite of the fact that some may be spurious, though it is notable that in many cases the extension is said to be explicitly to ‘bottom of a vessel,’ so that there are clearly also genuine cases. The association with ‘bottom’ or ‘base’ is attested by colexification in Buli, Efik, Hausa, Rendille, Buin, Burarra, Kwoma, Mali, Muna, Ngaanyatjarra, Kosarek Yale, Badaga (by an archaic term), Basque, Sora, Lesser Antillean Creole French, San Lucas Quiaviní Zapotec, Aymara, Carib, Guaraní, Ancash Quechua (where the relevant term is also an insult), Great Andamanese, Kapingamarangi, Lenakel, Malagasy, Rotuman, and Bislama. Tetun has *kidun-tahan* ‘bottom-flap,’ and Meyah *oskú ofogú* ‘bottom flesh’ (a semianalyzable term with ‘flesh’ is also found in Kyaka). Note also Kosarek Yale *di kumkum* ‘faeces neck/base’ and *dibomaak* ‘faeces-front.end/lower.part/bottom;’ semianalyzable terms with ‘bottom’ are furthermore found in Sko (where the putative constituent means “bottom of a four-legged animal” more specifically), Waris, and Wappo. Another cross-linguistic association related to a grammaticalization path (Heine and Kuteva 2002: 62) is that with ‘behind.’ Here, the same cautionary remarks made above for ‘bottom’ apply. ‘Back’ and/or ‘behind’ or ‘hinder part’ are colexified in Buli, Ngambay, Dongolese Nubian, Rendille, Badaga, Basque, Sora, Oneida, Carib, Huambisa, Rama, Fijian, Hawaiian (where the relevant term bears other related meanings as well), and Rotuman. Moreover, Chickasaw has *im-qshaka* ‘DAT-back,’ and Nuuchahnulth *?amas?akʔi*, analyzable as */ʔam-as-ʔakʔi/* ‘in.the.centre-approach-behind’ (a complex term featuring an element meaning ‘behind’ may have been present in Middle-Eastern Tasmanian). Somewhat similarly, ‘buttocks’ betrays an association with ‘end’ by colexification in Efik, Dongolese Nubian, Basque, San Lucas Quiaviní Zapotec, and Jarawara, while Abzakh Adyghe has *čʔentʔaʔ* */čʔe-n-tʔ(e)-ʔʔ(e)/* ‘end-RELAT/EPEN-poke.out-peak’ and Kosarek Yale the aforementioned term *di-bomaak* ‘faeces-front.end/lower.part/bottom.’

There are also contiguity-based associations with adjacent parts of the body: Dongolese Nubian, Rotokas, Toaripi, Badaga, Sora, Cheyenne, Lakhota, Pipil, Aguaruna, Bora, Jarawara, Hawaiian (the relevant term *ʔōkole* may be derived from *kole* ‘raw, inflamed, red,’ there is another Hawaiian term with a similar structure containing *kole*), and Bislama colexify ‘buttocks’ with ‘anus,’ and Baruya features a complex term with the constituents

'anus' and 'hill.' 'Buttocks' is colexified with 'hip' or 'haunch' in Dongolese Nubian, Kwoma, Kosarek Yale, Japanese, Kildin Saami (by the analyzable term *čuar-piell* 'pelvic-half'), Welsh, Upper Chehalis (by a term containing a constituent meaning 'elbow'; the gloss 'hips' is put in parentheses in the source), Yuki, Central Yup'ik (Nunivak Island dialect), Cashinahua, Piro, Hawaiian, Bwe Karen (where the relevant term contains *-ká* 'hind part'), and Sedang, in which latter the relevant term *kotei* is also the name of a gray bird. Japanese further has the analyzable term *den-bu* 'hip-part,' Bororo *etawu oto* 'hip point,' and Mandarin *pi4-gu* 'fart-upper.thigh/hip/member.' Mandarin is not the only language betraying an association with 'thigh': Ngambay and Kildin Saami (by the same analyzable term just mentioned) colexify 'thigh' with 'buttocks,' and Chickasaw has *iy-yobi' ishto-ka'* 'thigh be.big-NMLZ.' Toaripi, Kosarek Yale, Khalkha, Biloxi, Cheyenne, Kiowa, Lesser Antillean Creole French, Pawnee, Embera (by the term *āndáuboregáa*, analyzable as /āndáuboregáa/ 'rear/hind-fat'), Miskito, Wayampi, and Fijian colexify 'buttocks' with 'rump.' There is also one association with another body-part that seems metaphorical in nature, namely that with 'cheek.' Basque and Pawnee colexifies these meanings, while Muna has *bhaga-no koro* 'cheek-POSS buttock,' Kashaya *sili qapá* 'rear jowls,' and Wichí *towejch'alus* contains *wej* 'queue' and *ch'alu* 'cheek.'

As discussed above, Kosarek Yale has a complex term where one constituent is 'faeces;' this association is also found by colexification in Takia and by overt terms also in Bora (*námehéyu* /name-héyu/ 'faeces-CL.hole'), Jarawara (*joto-hoti* 'faeces-hole'), and Lenakel *nimwa-nisii-* 'covering-excrement-;' for the association with 'hole,' note also Kolyma Yukaghir *ńerčad-aqil* 'bad-hole.'

There are also metaphorical extensions of 'buttocks,' sometimes to more abstract meanings: Yoruba and Mali colexify it with 'reason,' Hausa and Yoruba with 'root' (Hausa also with 'foundation'), and Haida, Guaraní, Miskito, and Hawaiian with 'stern of boat.' Moreover, Rendille, Basque, and Kolyma Yukaghir colexify 'buttocks' with 'tail,' while Rama has *túkákás*, analyzable as /tuk-kás/ 'tail/end meat,' White Hmong *pob-tw* 'ball-tail,' and in Khoekhoe, the same root is used with alternating nominal designants to convey the two meanings. Similarly, Ngaanyatjarra colexifies 'tail of insect' specifically. Finally, Buli, Basque, Rama, Samoan, and Yay colexifies 'buttocks' with 'seat' (Samoan also with 'dwelling, residence' and 'station;' the meaning 'buttocks' is polite). Similarly, Yuki and Samoan feature terms derived from verbs meaning 'to sit.'

Other associations include: Buli colexifies 'buttocks' also with "outside, outer surface," as said of pots, and, by a different term, with 'descendants,' in particular 'grand-children.' Ngambay *bàgìrì* also denotes a "sieve, type of basket to remove husk from sesame, millet and gourd seeds." Yoruba colexifies 'buttocks' with 'waist.' Buin *koku* also means 'valley' and 'point of a banana rope,' while Muna *koro* also denotes a type of rooster, and One you also means 'call.' Abzakh Adyghe *č'ent'əʔ* also means 'kidneys' and Badaga *kuṇḍe* also 'olive' inter alia. Sora colexifies 'buttocks' with 'rectum,' and Chickasaw *ishkish* is also used to refer to the sexual organs in general. The Haida term *ruda* is also used with the meanings "the outside part of certain body parts," "that part of an island nearest to other land" and "side of house." Kiliwa *?uuw-h-pa?* is analyzable as 'vagina-3-set.round.object.down.' The Kiowa root *fei-* in *fei-dl* 'buttocks' also means 'calf' when used

with the noun postfix *-p*. Oneida *ohnátsha?* also means ‘small stool,’ Tuscarora *uhnèneh* and Guaraní *topyta* also ‘trunk’ (San Lucas Quiaviní Zapotec colexifies ‘trunk of car’ specifically), and the Wintu term *werem tah* contains *tah* ‘nearby.’ Miskito *pnata* also means ‘croup.’ Great Andamanese *ardama* appears to contain *dama* ‘lean’ – in contrast, Hani *daoqmeil* seems to contain *meil*, meaning ‘fat, obese’ inter alia. Hawaiian *lemu* is also used with the meaning ‘slow-moving, sluggish, lagging,’ and *ōkole* is also the name of a sea creature. Rotuman *muri* also means “young, not having reached maturity.”

### 107. The Calf

Representation: 79%

Motivated: 43.4%

Thereof Analyzable: 34.0%

Thereof Colexifying: 10.5%

Thereof by Contiguity: 13.0%

Thereof by Similarity: 28.5%

Recurrent associated meanings: leg/foot, belly/stomach, muscle, shin, egg, knee, pregnant/pregnancy, cheek, scrotum, fruit, flesh

Motivated terms for the ‘calf’ are more often morphologically complex than colexifying, and more often driven by metaphor than by metonymy. In most languages with complex terms, one of the constituents means either ‘leg’ and/or ‘foot.’ In ten sampled languages, ‘belly’ or ‘stomach’ is the meaning of the second constituent, as in Wichí *tot-kolo-ts’e* ‘POSS.INDEF-leg-paunch.’ Such terms are also attested in Efik, Sko, Toba, Malagasy, Mandarin, White Hmong, and Tetun. In addition Welsh colexifies ‘calf’ with ‘belly, womb’ and by another term with ‘belly, stomach’ (and has additional redundant complex terms), and Cashinahua has *bipustu*, where *pustu* is ‘stomach’ and *bi* ‘mosquito’ and ‘anteater,’ alongside *bi-tuxtu*, where *tuxtu* is ‘rounding.’ In Pawnee, furthermore, ‘calf’ is *kaac-karaar-u?* ‘gray-belly-NOM.’ Compare Sadovszky (1973) for discussion of this association and Matisoff (2004: 358) for further examples; Sadovszky also demonstrates a widespread association of ‘fish’ or more specifically ‘fish roe’ with the ‘leg’ and the ‘calf of the leg’ in Eurasia more specifically.

Similarly, Wayampi has *timã-pulu?a* ‘leg-be.pregnant,’ and this metaphorical transfer is realized by colexification with ‘visibly pregnant’ in Ngaanyatjarra, where the latter meaning is register-specific. Swahili has *shavu la mguu* ‘cheek of leg,’ and Hawaiian colexifies ‘calf’ with ‘cheek’ directly, alongside other meanings such as a container made from a long gourd. Biloxi has *yupkě’-i’iti’* ‘leg-egg.’ An analyzable term of the lexical kind is also found in Haida, and one of the derived kind in Burarra. In Takia, the ‘calf’ is *ŋe-n labe-n* ‘leg-3SG scrotum-3SG,’ while Hawaiian colexifies ‘calf’ and ‘scrotum’ (among other meanings). Yaqui has *woktomam* /wokim-tomam/ ‘feet-muscles,’ such a term (with ‘leg’ rather than ‘feet’) is also featured in Kashaya (here, *ʔahp<sup>h</sup>et* curiously colexifies ‘muscle’ with ‘mussel,’ in particular *Mytilus californianus*), while Khalkha, Cheyenne and Tehuelche colexify ‘calf’ with ‘muscle’ generally; moreover, Badaga *monne kaṇḍa* is analyzable as ‘lower muscle.’ There are also many other complex terms where one of the constituents is ‘foot’ and/or ‘leg’: Yoruba has *isu-ẹsẹ* ‘yam-leg,’ Basque *zango-sagar* ‘leg-apple,’ Upper Chehalis *sáwtiyq* contains *ʔáw-* ‘behind’ and *=iyq* ‘foot/leg,’ Chickasaw has *iyintakaali*, analyza-



ble as /iyyi' im-takaali-'/'leg DAT-be.hung.up.on-NMLZ,' San Mateo del Mar Huave *mi-pemb oleqjaran* 'AL.POSS-water.pitcher/gourd foot/leg,' Kiliwa *miy=ha?* 'leg-face' (marked with a question mark in the consulted source), Lesser Antillean Creole French *gwat janm* 'grater leg,' Santiago Mexquititlan Otomí *dot'uwa*, questionably analyzable as /tot'i-wa/ 'fold/bend-foot/leg,' Yana *ʒaʔpɔlgadu /ʒaʔpɔl-gaadu/* 'full/be.filled-leg,' Carib *-isai-punu* 'leg-chair,' Guaraní *tetyma ro'o /tetyma to'o/* 'leg flesh' (for which compare Kwoma *hapa omu* 'bone flesh/seed/fruit,' for which in turn compare Bwe Karen *kha-dě-θe* 'leg-narrowest.part-fruit'), Rama *kát-aring* 'foot-shit,' Wayampi *timākuʔa* 'leg-middle,' Fijian *temo ni yava* 'thick.part.of.limb POSS leg,' Rotuman *parpar ne lā /parpara ne lā/* 'soft ART.PL foot/leg,' Samoan *ate-vae* 'liver-foot/leg,' Vietnamese *bắp chân* 'shaft leg,' and Yay *ʔaay' ka' (raay')* 'goiter leg (mark).' Finally, Comanche has *taʔwiitsa*, seemingly analyzable as /ta-wiitsa/ 'foot-leg,' there is a semianalyzable term where the identifiable constituent means 'leg' in Bakueri, where it means 'back side of leg' specifically in Cayapa, and a semianalyzable term involving a lexical affix for 'foot, toe, leg from knee down' in Upper Chehalis.

Greek, Quileute, Cubeo, Miskito, and Imbabura Quechua colexify 'leg' with 'calf' directly (and Hani with 'lower leg below knee' more specifically), while the referent of the relevant Pipil term varies semantically between 'leg' in the Cuisnahuat dialect and 'calf of leg' in the Santo Domingo de Guzmán dialect. Semianalyzable terms are furthermore found in Bakueri, Buli, Mbum, Angkor (the meaning of the entire term is considered unsure by lexicographers; the unknown element is *mbosimundi*, compare *mbusimondi* 'middle sibling, thorax of spider'?), Kemtuik, Ngaanyatjarra, Sentani, Abzakh Adyghe, Badaga, Sora, Chickasaw, Highland Chontal, Oneida, Embera, and Lenakel.

Further, Badaga, Chickasaw, and Sedang colexify 'calf' with 'shin' (Sedang also with 'to pick vegetables,' 'to transplant rice'); analogously, Japanese has *fukura-hagi* 'swell-shin,' and Great Andamanese has *abchâltadama* containing *abchâlta* 'shin' and *dama* 'lean' and *tâl'ârdama* containing *tâ* 'bone' alongside the element meaning 'lean.' A semianalyzable term with 'shin' is also found in Carrier. Finally, Kosarek Yale and Badaga colexify 'calf' with 'knee,' Aguaruna with 'hollow of the knee,' and Yir Yoront has *nhal-kar* 'back/inside.of.knee-like.'

Other associations include: Hausa *sha ra'ba* is analyzable as 'drink dew' (with *ra'ba* however also meaning 'to crouch near something' *inter alia*). The whole term can also refer to a "metal ornamental point to a sword or knife sheath." The Khoekhoe root *lkhoe* yields the meanings 'calf (muscle)' and "meat of shank, knuckle" depending on the nominal designant suffixed. Muna *lambi*, as a verb, means 'to hang,' and Kosarek Yale *buding* is also the name of a tree frog species and a variety of taro. Basque (Lower Navarrese dialect) *aztal* also means 'heel,' and Kolyma Yukaghir *čĩnčade* (containing *čĩnčə* 'leg muscles') may also refer to the 'back part of boots.' The Kiowa root *tei-* in *tei-p* 'calf' also yields the meaning 'buttocks' when used with the noun postfix *-dl*, and Wintu *čuhcir* may be related to *čoy* 'to sprawl, stretch.' Xicotepec de Juárez Totonac *i'xchāpa'ta'kāt* contains *chāpá'* 'pole,' and Central Yup'ik *nakacugnaq* consists of *nakacuk* 'bladder' and the postbase (see § 4.4.2.) *-naq* 'one like,' while Copainalá Zoque *po'cpo'c* is a reduplication of *po'c* 'knot, trunk of tree, blain.' Chayahuita *pa'o ~ pa'huě* also means 'butt of a gun,' and Yanomámi *shiāpi* also means

‘thick part of an axe that doesn’t cut.’ For Ancash Quechua *pinkuullu*, compare *pinkullu* ‘slim flute.’ Tetun *kloor* may also refer to a ‘hamstring, hock,’ ‘footstring,’ or a ‘trace.’

#### 108. *The Cheek*

Representation: 93%

Motivated: 26.3%

Thereof Analyzable: 10.2%

Thereof Colexifying: 16.5%

Thereof by Contiguity: 20.3%

Thereof by Similarity: 4.2%

Recurrent associated meanings: face, jaw, mouth, cover, temple, chin, buttocks, edge, meat/flesh

‘Cheek’ (ignoring additional glosses like ‘side of face’) is often colexified with ‘jaw,’ namely in Buli, Yoruba, Gurindji, Yir Yoront (by the term *wal-kur* ‘cheek/temple-slashing.sword’), Welsh, Lesser Antillean Creole French, Wintu, Abipón, Miskito, Kapingamarangi (by the term *gau wae*, which is seemingly analyzable as ‘side foot/leg’), and Manange (see Buck 1949: 220–221 for the strong diachronic ties between terms for ‘cheek,’ ‘jaw’ and ‘chin’ in Indo-European). In Basque there is dialectal variation of the meaning of *baraila*, meaning ‘jaw’ alongside ‘revelry, fray’ in the Biscay dialect and ‘cheek’ in the Zuberoan dialect; further, for Hani *baqba*, compare *baqxoq* ‘jaw.’ Common are also associations with ‘face’ (and sometimes due to further colexification also with ‘front’), by colexification in North-eastern Tasmanian, Abzakh Adyghe (by the analyzable terms *nek<sup>o</sup>ʔəs<sup>e</sup>*, containing *ne* ‘mouth’ and *s<sup>e</sup>* ‘skin’ and *nek<sup>o</sup>ʔəs<sup>he</sup>*, containing *ne* ‘mouth’ and *s<sup>he</sup>* ‘upper part’), Chickasaw, Lesser Antillean Creole French (where also ‘figure, form’ and other meanings are colexified), Nez Perce, Wintu, Aguaruna, Bororo, Carib, Hupda, Ancash Quechua, Tehuelche, and Sedang. In addition, Abzakh Adyghe has *nepas<sup>e</sup>* /nape-ś<sup>e</sup>/ ‘face-skin’ (a semianalyzable term with ‘skin’ is found also in One), Kiowa has *ʔou-ḡḡ<sup>egyh</sup>* ‘face-in.the.middle.of,’ Guaraní *tova-yke* ‘face-side,’ Imbabura Quechua *ñawi chichu* ‘face pregnant,’ and Rama *mngut kás ~ múngut kás*, containing *ngút* ‘face’ and *kás* ‘meat.’ A term with identical structure is found in Maxakalí, and for Carrier, the source remarks: “the Carrier for cheek seems to mean face-profile, or nearly so” [sic!]. Semianalyzable terms for ‘cheek’ where the identifiable constituent is ‘face’ are furthermore found in Xicotepec de Juárez Totonac, Yana, and Tsafiki. Alongside the Abzakh Adyghe term just mentioned, there are also associations between ‘cheek’ and ‘mouth’ in other languages (in diachrony, Romance terms for ‘mouth,’ such as Spanish *boca*, go back to Latin *bucca* ‘cheek,’ see also Buck 1949: 221 on the cognacy of Germanic terms for ‘cheek’ with Avestan ‘mouth’): they are colexified in Burarra (also with ‘lips,’ the term has broad reference to the ‘mouth area’ in general), and analyzable terms are found in Mbum (*fã-háù* ‘sack-mouth;’ similarly, ‘bag’ and ‘cheek’ are colexified in Ngambay), Sora (*kub-ma<sup>b</sup>-tam-ən* ‘unite/be.covered-??-mouth-N.SFX;’ similarly, Efik has a derived term from a verb meaning ‘to cover’ and Rotuman colexifies ‘cheek’ and ‘(to) cover’ inter alia), and Kiliwa (*yuw=ha?=kw-cas* ‘eye=mouth=WH-??,’ with *yuw=ha?* meaning ‘face;’ there is a semianalyzable term where the identifiable constituent means ‘eye’ in Abzakh Adyghe). Dialectally, the meaning of Yana *bal(la)* varies between ‘cheek’ and ‘mouth.’ Semianalyzable terms with ‘mouth’ are

also featured in Baruya and Kiliwa. Dongolese Nubian and Yir Yoront colexify ‘cheek’ with ‘temple’ (Yir Yoront alongside ‘ear’ and ‘sleep’), and Burarra and perhaps Abipón with ‘chin.’ Basque and Pawnee colexify ‘cheek’ with ‘buttocks,’ Katcha colexifies ‘cheek’ with ‘edge,’ and Copainalá Zoque has *aca-pac* ‘edge-bone’ (an element meaning ‘bone’ is also etymologically recoverable in the Ket term for ‘cheek’).

Other associations include: Lavukaleve *hou* is also used with the meaning ‘smoke,’ while Muna *bhaga* also means ‘molar tooth’ (there is a semianalyzable term containing an element meaning ‘tooth’ in Berik; compare also Nunggubuyu *ramara* ‘cheek’ and *ra:* ‘tooth’). Toaripi *heva* also means ‘gill.’ *Sko ðebi* appears to be analyzable as /ðe-bí/ ‘penis-shell/floor’ and also means ‘testicle,’ while Tuscarora *unhú?weh* also means ‘areola of nipple.’ Arabela *sacomara* ‘cheek pouch’ appears to consist of *saco* ‘hollow fruit’ and *mara* ‘swamp.’ Jarawara *abate/ebete* also means ‘tongue,’ and Tehuelche *q’ape-n-k’en* is analyzable as ‘be.red-??-LOC.’ Fijian *balu* also denotes the ‘sides of the head of a club,’ while Hani *baqba* may be related to *baq* ‘direction,’ ‘thin’ and *ba* ‘light in color, white.’ Hawaiian colexifies ‘cheek’ and ‘calf of leg,’ *papālina* ~ *pāpālina* is perhaps analyzable as /papa-lina/ ‘flat.surface-soft’ (both *papā* and *pāpā* exist as well, but do not have meanings standing in any obvious relationship with ‘cheek’), and Bwe Karen *-bɔ* also means ‘packet, parcel’ and “creeper with a bitter shoot.”

#### 109. *The Chin*

Representation: 90%

Motivated: 44.2%

Thereof Analyzable: 12.8%

Thereof Colexifying: 33.0%

Thereof by Contiguity: 40.5%

Thereof by Similarity: 3.8%

Recurrent associated meanings: jaw, beard, mouth, lower/below, bone/skeleton, cheek

‘Chin’ is very commonly colexified with ‘(lower) jaw’ cross-linguistically (see Buck 1949: 220–221 for the strong diachronic ties between terms for ‘cheek,’ ‘jaw,’ and ‘chin’ in Indo-European). This is found in as many as 44 languages, namely Yoruba, Baruya (by the analyzable term *maanaginya* /maanga-yaginya/ ‘mouth-bone’), Kaluli, Kwoma (by the term *teekibi*, perhaps containing *teeki* ‘to string, tense’ and *bi* ‘point’), Kyaka, Lavukaleve, Mali, One, Rotokas, Southeastern Tasmanian, Toaripi, Kosarek Yale, Abzakh Adyghe (by the analyzable term *ze-pqʔ* ‘mouth-skeleton’), Greek, Japanese, Ket (by a semianalyzable term containing *it* ‘tooth,’ an element with that meaning is also found in the Berik, Embera and Kaingang terms), Khalkha, Nivkh, Welsh, Kolyma Yukaghir, Blackfoot, Upper Chehalis, Cheyenne, idiolectally in Chickasaw, Itzaj, Kiliwa (by a semianalyzable term containing *ha?* ‘mouth’), Nez Perce, Nuuchahnulth (by the analyzable term *hiiniilaksut* /hiiniilā-ʔaksut/ ‘lower/below-at.the.mouth/at.the.lips’ which also denotes the ‘lower lip’), Pipil (by the analyzable term *-te:n-tsi:ka-w* ‘-mouth-ant-POSS’), Arabela, Aymara, Cashinahua, Embera (where the meanings are associated with different genders), Guaraní, Jarawara, Macaguán (by a semianalyzable term containing *-buk* ‘heel’), Rama (by a semianalyzable term containing an element meaning ‘tree, foot’ and colexifying ‘grater’), Tehuelche, Yanomámi, Mandarin, White Hmong, Samoan (by the analyzable term *‘au-vae*

‘CLASS.long.or.narrow.things-foot/leg’ which also means ‘foot of hill), Takia (by the analyzable term *awa-n to-n* ‘mouth-3SG arm-3SG’), Tetun, and Yay; a semianalyzable term is found in Mandarin. Complex terms for ‘chin’ on the basis of ‘jaw’ are Biloxi *yatka’ psû”ti’ ~ i’yatka’ psû”ti’ ~ nyatka’ psû”ti’* /*yatka’ pû”tsa/* ‘jaw sharp,’ Cayapa *tejmashmutu* /*tejmashi-mutu/* ‘jaw-corner,’ and Chayahuita *hui’nin cohuiatē* ‘child/prow jaw.’ Associations with ‘beard’ are also quite common, also in Indo-European (Buck 1949: 224). ‘Beard’ and ‘chin’ are colexified in Dongolese Nubian (which also colexifies ‘whiskers’), Gurindji, Ngaanyatjarra, Tasmanian (Northeastern, Middle Eastern, and Western), Badaga, Embera, Fijian, Rotuman, and Tetun. Further, Kanuri has *njiti-rám* ‘beard-place.of,’ in Swahili, the term for ‘chin’ consists of that for ‘beard’ and a noun class prefix, and semianalyzable terms are found in Buli and Sentani (note that Ancient Greek had a term literally translated by Buck 1949: 224 as ‘that which has a beard on it’ and Russian one literally meaning ‘what is under the beard,’ there is a semianalyzable term where the identifiable constituent means ‘under’ inter alia in Wintu). Bora has *újcawa*, with the same root possibly also occurring in the word for ‘beard,’ but segmentation is uncertain. As already seen in some of the above terms, there are also associations between ‘chin’ and ‘mouth.’ Burarra has a general term referring to the mouth area in general, including the ‘chin,’ and complex terms are found in Baruya (*maanaginya* /*maanga-yaginya/* ‘mouth-bone’), Abzakh Adyghe (*žepq’ /žepq’(ə)/* ‘mouth-skeleton’), Sora (*də’nanjul’tam- ~ dənankultam-*, containing *dənanj-* meaning ‘obstruction, bar’ inter alia and *tam* ‘mouth’), Kiowa (*beidl-t’ēj’m* ‘lip/mouth-bone’), Nuuchahnulth (*hiiniiiłaksuł /hiiniiiłaxuł-* ‘lower/below-at.the.mouth/at.the.lips’), Pipil (*te:n-tsi:ka-w* ‘mouth-ant-poss’), Lengua (*koning atang* ‘below mouth’), and Mandarin (*xia4-ba*, with *xia4* meaning ‘down, lower’ and *ba* referring to the ‘mouth’). Semianalyzable terms are found in Nunggubuyu and Kiliwa. For the associations with ‘bone’ above, note also that Carrier has a semianalyzable term where the identifiable constituent bears that meaning. Finally, Burarra and perhaps Abipón colexify ‘chin’ with ‘cheek.’

Other associations include: Dadibi *penani* also means ‘adultery, fornication,’ One sesu also ‘to squish,’ Kosarek Yale colexifies ‘chin’ with “wattle of an agama,” and Basque *kokots* also means ‘snout,’ ‘calyx,’ and ‘mesh.’ Carrier *-yēta* contains *ta* ‘surface,’ Central Yup’ik *tamlu ~ tamluq tamu-* ‘to chew once,’ and Jarawara *enekiri/enekiri* also means ‘gill.’ Macagúan *pipumáchipla* also can refer to the ‘forehead,’ Toba *lqa’* also is glossed as Spanish ‘pera,’ which is either ‘pear,’ ‘bulb,’ or ‘signal horn’ in English, while Wayampi *eni-wa* is analyzable as ‘saliva-eater.’ Fijian *kumi* also denotes ‘Tongan cloth,’ and Bwe Karen *khe* also means ‘to be bitter’ inter alia. Hawaiian *auwae* inter alia also denotes the “curved notch cut on the outer side of a post below the base of a tenon.” Rotuman *kumkumu* also denotes a species of crab, and Tetun also means ‘ambush’ and “notch in wooden columns of buildings for the positioning of beams.”

#### 6.2.2.110. The Eyeball

Representation: 36%

Motivated: 86.9%

Thereof Analyzable: 81%

Thereof by Contiguity: 17.5%

Thereof Colexifying: 7.5%

Thereof by Similarity: 65.4%

Recurrent associated meanings: eye, seed, egg, pupil, child, fruit, grain, stone, ball, round object

Terms for ‘eyeball’ are overwhelmingly morphologically complex in the languages of the sample, and are often metaphor-based. Typically and obviously, ‘eye’ is one of the constituents, and the second one in metaphor-driven terms typically denotes a small round object. Ngaanyatjarra has *kuru yurniny(pa)* ‘eye seed/grain,’ and terms with an association with ‘seed’ are also featured in Buli, Efik (where ‘seed’ is *mpa’sip* ~ *mkpasip* /mp’kō-sip/ ‘thing-be.small’), Kwoma, Muna, Sahu, Kosarek Yale, Sora, Itzaj, Lesser Antillean Creole French, Wintu, Guaraní (though note that *ta’yi* ‘seed’ is a diminutive of *ta’y* ‘son, clot’), Miskito, Bwe Karen, and Sedang, while Yay features a semianalyzable term where the identifiable constituent is ‘seed.’ Kosarek Yale *wana* colexifies ‘seed’ with ‘eye,’ ‘heart,’ ‘egg,’ ‘fruit,’ and ‘child’ (compare § 6.2.3.1.), and, by virtue of this, the term for ‘eyeball’ *heing wana* also betrays an association with these meanings. An association with ‘egg’ is also found in a number of other sampled languages, as in Toba *l-’ai l-co’oue* ‘3SG.POSS-eye 3SG.POSS-egg,’ and also in Efik, Koyraboro Senni, Mbum, Yoruba, Sora, Wintu, Fijian, Kapingamarangi, and Lenakel. Wappo has *huči:lél* /huči-lél/ ‘eye-stone’ (which also denotes ‘hailstone’), and analogous complex terms without the additional colexified meaning exist in Koyraboro Senni and Kashaya (in Koyraboro Senni, the constituent meaning ‘stone,’ however, is said to assume the meaning ‘egg’ in compounds as well). Similarly, Kosarek Yale, Kwoma, and Miskito colexify ‘seed’ and ‘fruit,’ and therefore, in these languages, ‘fruit’ is (also) the source concept for the metaphorical transfer; the same situation obtains in Tetun. Further, Piro colexifies ‘large round fruit’ with ‘ball,’ and analogously to English *eyeball*, ‘ball’ is the meaning of the second constituent in Lake Miwok (*šút polólo* ‘eye ball’), in Kiliwa, and, by virtue of the language’s pattern of colexification, in Piro. Similarly, Basque has *begi-globo* ‘eye-globe.’ Katcha has *bibala m-æ* ‘child ??-eye,’ and such terms are also found in Kosarek Yale, by virtue of the large semantic range of the constituent *wana* as discussed above, Welsh (colexifying ‘boy, son’ specifically rather than ‘child’ generally; this term is marked as being obsolete), Samoan, and, because of the internal structure of the term for ‘seed’ discussed above, also in Guaraní. A semianalyzable term with ‘child’ is featured in Yay, and similarly, Carrier has *tšûten-pê-šta* ‘child-where-with-sit.in.’ This pattern is likely intertwined with the association between the ‘eyeball’ and the ‘pupil’ (see also section 130); ‘eyeball’ and ‘pupil’ are colexified (sometimes by analyzable terms discussed elsewhere in this paragraph) in Hausa, Katcha, Yoruba, Welsh, Wintu (where the relevant term also denotes the ‘iris’ at the same time), and Guaraní.

Complex terms with ‘round object’ alongside ‘eye’ are featured in Chickasaw (*ishkin lobo* ‘eye round.and.firm.object’) and Oneida, and Blackfoot has *ohkomapinssin* /ohkom-aapini-hsiN/ ‘be.round-eye-NMLZ.’ There is a complex term with ‘white’ in Kaingang (*kanē kupri* ‘eye white’). Other complex terms of the lexical type with ‘eye’ being one of the constituents are Hausa *k’wayar ido* ‘grain eye’ (this association is also present in Ngaanyatjarra, as noted above, as well as Lesser Antillean Creole French. Similarly, Hani *miavneev* ‘eyeball, eye’ is derived from *miav* ‘eye’ by means of *neev*, the classifier for beans and grains), Mbum *máà-yâr* ‘mother-eye,’ Ngambay *dò-kèm* ‘head/on-eye,’ Kosarek Yale

*heing ngei* ‘eye ground/basis,’ Sora *‘ran-a-‘mad-ən* ‘gold-eye-N.SFX’ (glossed in the source as ‘gem of the eye’ and marked with a question mark), Welsh *cannwyll y llygad* ‘candle DET eye,’ Cheyenne *menoo’éxané*, containing *mené* ‘berry’ and *éxané* ‘eye’ (gloss in the source: “globular (that is, berry).shaped-eye”), Lesser Antillean Creole French *koko zîé* ‘coconut eye,’ Jarawara *noki kori/noko korone* ‘nakedness eye,’ Samoan *‘i‘oi-mata* ‘tuber/corm-eye,’ and Bislama *mabol blong ae* and *pi blong ae* ‘marble POSS eye.’ Moreover, ‘eye’ and ‘eyeball’ are colexified directly in Buli, Koyraboro Senni, Toaripi, Yir Yoront, Laz, Nivkh, Jarawara, and Hani. In this case, the ‘eye’ as a physical object is probably the colexified meaning more precisely; semianalyzable terms for the ‘eyeball’ on the basis of ‘eye’ are found in Koyraboro Senni (where the second constituent is diachronically relatable to a word meaning ‘belly’), Yir Yoront (where the second constituent diachronically goes back to a word meaning ‘moon’), and Bislama.

Other associations are few in number: the Burarra term *munbarra* consists of the noun class prefix *mun-* and *barra* ‘rear end,’ Greek colexifies ‘bulb’ with ‘eyeball,’ and Nuuchahnulth *łiskčirnim* contains the verb root *łisk-* ‘for the eyes to flash white.’ San Lucas Quiaviní Zapotec *bàa’ah* is also used with the meaning ‘iris of the eye,’ and Piro *tskata* also means ‘cliff’ or ‘bank.’ Hawaiian *pipi* also denotes the ‘Hawaiian Pearl Oyster’ inter alia, and for *ōnōhi*, compare *nōhi* ‘bright-colored, vivid’ and the ‘similitude’ prefix *ō-*. This term also means ‘center, setting (as of a ring),’ and *pona* also means ‘socket,’ among other meanings.

### 111. *The Eyebrow*

Representation: 92%

Motivated: 44.6%

Thereof Analyzable: 35.3%

Thereof Colexifying: 9.3%

Thereof by Contiguity: 15.7%

Thereof by Similarity: 9.8%

Recurrent associated meanings: eye, hair, eyelash, fur, feather, forehead, edge/fringe, wool, bone, brow of hill, eyeridge, joint

Motivated terms for ‘eyebrow’ are overwhelmingly of the lexical type, with ‘eye’ typically being one of the constituents. Unsurprisingly, the meaning of the other constituent is often ‘(body) hair,’ as in Highland Chontal *libimi gahu* /*libimi lahu*/ ‘body.hair eye.’ Such terms are also found in Kanuri, Koyraboro Senni, Ngambay, Berik, Dadibi, Kwoma, Toaripi, Badaga, Sora, Chickasaw, San Mateo del Mar Huave, Itzaj, Yuki, Cayapa, Jarawara, Miskito, Piro, Ancash Quechua, Rama, Bislama, Lenakel, White Hmong, and Yay. This pattern is also found in Tsafiki *ca’cá chidé fu* ‘eye tree/bone body.hair/fur/feather.’ In fact, an association with ‘bone’ is also attested in Ket and Pawnee (Pawnee, for instance, has *kirikiisu* /*kirikiisu-u*/ ‘eye-bone-NOM’). The Pawnee term also denotes the ‘superciliary ridge, supraorbital bone,’ and this may be the explanation for the association in the other languages as well (for instance, *ca’cá chidé* may be the term for the ‘ciliary’ in Tsafiki). Similarly, One has *namnaalo palla* ‘eyeridge body.hair’ (Buin directly colexifies ‘eyebrow’ and ‘eyeridge’ inter alia), Hawaiian has *ku’eku’e-maka* ‘joint/elbow/wristbone/knuckle-eye,’ and Tetun *matan-fukun* ‘eye-knot/joint/knuckle.’ There are also complex terms where the constituents are ‘eye’ and ‘fur’ (although it cannot be excluded that, by colexification of ‘fur’ and ‘body

hair,' these in fact could be subsumed under terms associated with '(body) hair'). Terms betraying an association with fur, either because this is the single meaning of the constituent as in Kolyma Yukaghir *and'an-pugulbe*: 'eye-fur,' or more frequently because 'fur' is colexified with 'body hair' are found in Buli, Badaga, Berik, Kwoma, Chickasaw, Kashaya, Rama, Tsafiki, Wayampi, Lenakel, Samoan, and Yay. In many of these languages, the relevant constituent also colexifies 'feather.' This association is present in Buli, Berik, Kwoma, Chickasaw, Highland Chontal, Rama, Tsafiki, Bislama, Lenakel, Yay, while Hani has *miav-hao* 'eye-feather.' Moreover, in Sora and Samoan, there is also an association with 'wool' due to colexifying structures in constituents of complex terms. Furthermore, Hupda has *kəwəg-pəw* 'eye-edge,' and, similarly, Cubeo *ēca-me* 'fringe-CLASS.LIKE.THREAD,' while Abzakh Adyghe directly colexifies 'eyebrow' and 'edge.' Colexification of 'eyebrow' with edges of specific objects is attested for Hausa and with 'edge of mountain' (as well as 'projection in cliff') in Khoekhoe (compare the parallel association between 'eyebrow' and 'mountain' evidenced by cognates in Irish, Latvian, and Albanian, Buck 1949: 220). Other complex terms where one of the constituents is 'eye' are: Efik *nditan-ēnyin*, seemingly analyzable as 'stupidity-eye,' Anggor *himboari-siri* 'eye-root/clan/origin,' Badaga *kañnu eme* 'eye lid/lash,' Basque *betgain* /begi-gain/ 'eye-above' (this may be calqued, given that French *sourcil* and Spanish *ceja* go back to Latin *super-cilium* 'above-eyelid'), Lake Miwok *šut límme* 'eye brush,' Lakota *ištáhe* /ištá-hé/ 'eye-mountain,' Arabela *namijia-qui* 'eye-CLASS.CLOTH,' Hupda *kəwəg b'ək cū'h* containing *kəwəg* 'eye' and *b'ək* 'skin,' Samoan *tuā-mata* 'beyond/across-eye,' and Takia *mala-n ddawe-n* 'eye-3SG handle-3SG.' There are semianalyzable terms in many languages: Bakueri, Koyraboro Senni, Kemtuik, Lavukaleve, Kosarek Yale, Abzakh Adyghe, Bezhta, Blackfoot, Carrier, Chickasaw, Lake Miwok, Nez Perce, Santiago Mexquititlan Otomí, Pipil (Santo Domingo de Guzmán dialect), Xicotepec de Juárez Totonac, Tuscarora, Wappo, Yana, Aymara, Bora, Macaguán, Hawaiian, Malagasy, Manange, Rotuman, and Sedang.

Several of the terms mentioned so far colexify 'eyebrow' with 'eyelash' (and there is semantic shift between these two meanings in Indo-European, Buck 1949: 219). This is the case in Buli, Koyraboro Senni, Berik, Kwoma, Kyaka, Sora, Chickasaw, Itzaj, Kashaya, Lake Miwok, Cayapa, Imbabura Quechua, Tehuelche, Bislama, Lenakel, Hani, Manange, White Hmong, and Yay. Furthermore, this pattern of colexification occurs in Burarra, Ngaanyatjarra, Nunggubuyu, Tasmanian (Middle-East and Southeast), Cheyenne, perhaps in Highland Chontal, Pipil, and Chayahuita by terms with no internal morphological structure.

However, there are also terms with constituents meaning 'hair' where the second element is not 'eye.' Thus, alongside a semianalyzable term in Cahuilla, Kyaka has *lembaki-sa emanji ~ lembaki-si yamanji* 'eyelid-LOC body.hair/fur,' San Lucas Quiaviní Zapotec *gyihch lahq ~ gyihch cyi'lahq* /gyihch cyi'lahq/ 'hair eyebrow.ridge,' Tehuelche *kašč'ex ~ kašč'ex*, containing *-aš* 'in, inside' and *č'ex ~ č'ex ~ č'e:x* 'body hair, wool' (alongside the similar term *lašč'ex ~ ač'ex ~ lačx ~ lač'ex ~ lašč'ex ~ a:č'ex ~ ač'ex*, which also means 'beard'), and Yanomámi *wəyomahiki kōi* 'superciliary.arch body.hair.' The Embera and Wayampi terms betray an association with 'forehead': *dráthu-kára* 'forehead-body.hair' and *apikā-l-a* 'forehead-of-hair' respectively, while 'forehead' and 'eyebrow' are colexified in Rama and

Sedang (see Buck 1949: 219 on diachronic connections in Indo-European). Complex terms where one of the constituents is ‘forehead’ are also found in Comanche (*kaʔibuh̥ ~ kaʔibuu*, analyzable as /kaʔi-puh̥/ ‘forehead-fuzz’) and Kiliwa (*miʔ=kw-m-puul=ny+miʔ* ‘forehead=OBJ-??-hat=POSS+forehead’) alongside semianalyzable terms in Biloxi and Carib. Imbabura Quechua has *ñawi milma* ‘face wool,’ and a semianalyzable term where the identifiable constituent means ‘face, eye’ is found in Yana, and one where it means ‘face, forehead’ in Upper Chehalis.

Other associations include: Rotokas *uvu keru* seems to contain *uvu* ‘sense something, hear’ and *keru* ‘harden like bone,’ while Sko *lúbi* is also used with the meaning ‘temple.’ Welsh *ael* also means ‘aisle’ and ‘litter.’ Cheyenne *véʔhenôse* is related to a verb meaning ‘be singed,’ and Haida *skʔyaayi ~ skʔyah ~ skʔy@s* is also used with the meaning ‘the first step down into a housepit.’ Arabela *susuque*, containing *-que* ‘cloth,’ also means ‘cloth with moths,’ and the relevant Cubeo term also denotes the ‘tilde’ in typography. Guaraní *tyvyta* appears to contain *tyvy*, which can mean either ‘younger brother’ and ‘grave,’ or *ty* ‘urine.’ Fijian *vacu* is also used to refer to “the place of insertion of the legs of a crab into its shell” and means ‘to punch with the fist’ as a verb, Kapingamarangi *himada* contains *mada* ‘to look, see,’ and Vietnamese *mày* is also the personal pronoun for the second person singular. Finally, Lesser Antillean Creole French *sousi* also means “care, solicitude, anxiety.”

#### 112. *The Eyelash*

Representation: 87%

Motivated: 61.4%

Thereof Analyzable: 53.6%

Thereof Colexifying: 8.0%

Thereof by Contiguity: 28.3%

Thereof by Similarity: 9.2%

Recurrent associated meanings: eye, hair, eyebrow, fur, feather, eyelid, wool, leaf, brush, beard

As with the ‘eyebrow’ discussed in section 111 terms of the lexical type, with constituents meaning ‘eye’ and ‘(body) hair’ are frequent for the ‘eyelash’ as well. Terms such as Hupda *kəwəg pãʔt* ‘eye hair’ are found in Bakueri, Efik, Koyraboro Senni, Berik, Dadibi (*gedu páde nisi* ‘eye near/almost hair,’ as opposed to *gedu nisi* ‘eyebrow’), Kwoma, Muna, Meyah, Toaripi, Kosarek Yale, Abzakh Adyghe (with an additional element present: *nebžac* /ne-bž(e)-c(e)/ ‘eye-horn-hair/fur/feather/wool’), Basque, Japanese, Ket, Sora, Carrier, Chickasaw, Highland Chontal, Itzaj, Kiowa, Pawnee (*kiriktaacpickiic*, the relevant term, contains *kirik-* ‘eye,’ *raac-* ‘pubic hair,’ and *kiiʔac* ‘long;’ it also denotes the ‘persimmon’), Santiago Mexquititlan Otomí, San Lucas Quiaviní Zapotec, Aguaruna, Bororo, Cavineña (*yatuca cuatsa tsaru*, containing *yatuca* ‘eye,’ *tsaru* ‘hair’ and presumably *e-cuatsa* ‘mouth’), Cayapa, Cubeo, Embera (*dau-i-kára* ‘eye-border-body.hair’), Maxakalí, Piro, Rama, Tsafiki, Wayampi, Yanomámi, Bislama, Great Andamanese, Hani, Lenakel, Malagasy, White Hmong, Sedang, Tetun (where the term is figuratively also used with the meaning “opportunity, chance”), and Yay. Very similar metaphor-driven terms are found in Lavukaleve (*lemi ohal* ‘eye leaf’), Samoan (*lau-mata* ‘leaf eye’), and Fijian (*bebekanimata* /bekabeka-ni-mata/ ‘coconut.leaves-POSS-eye;’ for this term, note also that ‘eyelash’ and “cocoa husk



[sic!]" are colexified in Rotokas). The association with 'leaf' is also present in Chickasaw, due to colexification of 'body hair,' 'fur,' and 'feather.' Note also Lake Miwok *šút límme* 'eye brush' and Bora *hállulli /hálluu-lli/* 'eye-CL.brush.' Katcha has *miziri m-ye* 'eyebrow ??-eye,' Arabela *namijia susuque* 'eye eyebrow,' and Vietnamese *lông mi* 'body.hair eyebrow' (an optional complex term of this kind is also found in Cheyenne, where 'eyebrow' and 'eyelash' are colexified). Other complex terms of the lexical type where 'eye' is a constituent are Mbum *tí-yâr* 'coming.out.from-eye,' Anggor, curiously, *himboari yapisendi* 'eye lid,' Badaga *kañnu hođe* 'eye bush/thicket/wood,' Greek *matóklado*, which is perhaps analyzable as /mát-o-klado/ 'eye-STEM.FORMATIVE-branch,' Kolyma Yukaghir *aŋd'a-šepil* 'eye-door' (denoting the 'upper eyelash' and 'upper eyelid' specifically), Cahuilla *puščávay /-puščávay-a/* 'eye-fall-?,' Kashaya *huʔu pitem?*, analyzable as /huʔuy pitem?/ 'eye droop.of.eyes,' Jarawara *noki masiri/noko masiri* 'eye grass/bird.species,' Miskito *won nakra taya* 'BODY.PART eye skin' (this term also denotes the 'eyelid'), and Wichí *tottefwis*, containing *tef* 'eyes' and *wis* 'larvae.' There are terms with 'fur' as the second element alongside 'eye' (which is frequently colexified with 'hair') in Buli, Berik, Kwoma, Kyaka, Meyah, Muna, Toaripi, Kosarek Yale, which also colexifies 'beard,' Abzakh Adyghe, Chickasaw, Kashaya, Lesser Antillean Creole French, Huambisa, Lenakel, Samoan, Tetun, and Yay, while Welsh has *blew yr amrant* 'hair/fur GEN eyelid'). In Buli, 'hair' and 'feather' are colexified. This is also the case in Berik, Kwoma, Meyah, Muna, Toaripi, Kosarek Yale, Abzakh Adyghe, Chickasaw, Huambisa, Tsafiki, Bislama, Lenakel, Samoan, Sedang, Tetun, and Yay, so that the relevant terms in these languages also bear an association with 'feather;' Hani has *miav-hao* 'eye-feather' for 'eyebrow' and 'eyelash,' without colexification of 'hair and feather.' Imbabura Quechua *ñawi milma* is analyzable as 'face wool,' and due to colexification with 'hair' and/or 'fur,' the association with 'wool' is also present in Abzakh Adyghe, Basque, Sora, Huambisa, and Samoan. There are semianalyzable terms involving a constituent 'eye' in Khoekhoe, Koyraboro Senni, Kosarek Yale, Bezhta, Khalkha, Blackfoot, Upper Chehalis, Chickasaw, Ineseño Chumash, Comanche, Haida (colexifying "ray of the sun shining through clouds"), Kiliwa, Lakhota, Tuscara, Wappo, Wintu, Yana (where 'eye' and 'face' are colexified, the reference of the term is considered dubious), Copainalá Zoque, Kaingang, Manange, and Takia.

As already seen in some of the terms just mentioned, associations between 'eyelash' and 'eyelid' are also common. They are colexified in Swahili, Gurindji, Badaga, Kolyma Yukaghir ('upper eyelash' and 'upper eyelid' specifically), Upper Chehalis, Santiago Mexquititlan Otomí, Cavineña (by analyzable terms containing 'eye' and 'hair' in both aforementioned languages), Miskito, and Hawaiian (by the term *lihilihi*, the reduplication base of which means 'edge,' colexifying also 'lace,' 'to crochet,' and denoting a variety of sweet yam). Similar to the Welsh term already mentioned but without colexification of 'hair' with 'fur,' terms containing elements meaning 'hair' and 'eyelid' are also attested in Yoruba and Kyaka. Similarly, Embera has *daúikará* which appears to be analyzable as /dauí-ikaráa/ 'eyelid-beard,' and Guaraní has *tope-a* 'eyelid-fruit.' San Lucas Quiavini Zapotec has *gyihch bàa'ah* 'hair eyeball' (this term also denotes a card game), Tehuelche *kašč'ex ~ kašč'ex*, containing *-ašč* 'in, inside' and *č'ex ~ č'ex ~ č'ex* 'body hair, wool' (there also is the similar term *lašč'ex ~ ašč'ex ~ lašč'ex ~ lašč'ex ~ lašč'ex ~ a:šč'ex ~ a:šč'ex* which also means

‘beard’), and semianalyzable terms with ‘hair’ are in addition present in Aguaruna, Carib, and Lengua. Several of the terms mentioned so far colexify ‘eyelash’ with ‘eyebrow’ (see Buck 1949: 219–220 for Indo-European evidence for this association). This is the case in Buli, Koyraboro Senni, Berik, Kwoma, Kyaka, Sora, Chickasaw, Itzaj, Kashaya, Lake Miwok, Cayapa, Imbabura Quechua, Tehuelche, Bislama, Lenakel, Hani, Manange, White Hmong, and Yay. Furthermore, the association is in addition found by unanalyzable terms in Burarra, Ngaanyatjarra, Nunggubuyu, Tasmanian (Middle-eastern and Southeastern), Cheyenne, Pipil, and Chayahuita, and it may be present in Highland Chontal, where the terms differ only in one segment which may be a typographical error.

Given that there are many languages without colexification of ‘eyebrow’ and ‘eyelash,’ which, however, still express both meanings with complex terms on the basis of ‘eye,’ it is interesting how ‘eyebrow’ and ‘eyelash’ are differentiated. One solution, adopted in Toaripi, is to use the lexical differentiation between different types of hair present in the language: *ofae ve mehe* ‘eye POSS hair’ is ‘eyelash,’ while *ovo-tui* ‘eye-hair.of.head’ is ‘eyebrow.’ Another option, present for instance in Badaga, is to have a metaphor-driven term for one of the meanings: *kaŋŋu hoe* ‘eye bush/thicket/wood’ is ‘eyelash,’ and *kaŋŋu mailu* ‘eye hair/fur’ is ‘eyebrow.’ Yet another one, present for instance in Basque, is to have a constituent neither meaning ‘hair’ nor being metaphorically related to it: *betile* /*begi-ile*/ ‘eye hair’ is ‘eyelash,’ and *betgain* /*begi-gain*/ ‘eye-above’ is ‘eyebrow.’ Moreover, Samoan has *fulu-mata* ‘fur/wool/feather-eye’ for ‘eyelash,’ and *fulufulumata*, with the word for ‘fur, wool, feather’ reduplicated, for ‘eyebrow.’ Finally, terms for both meanings may feature constituents meaning ‘eye’ and ‘hair,’ but one of them, typically ‘eyebrow,’ has another additional constituent. Thus, San Mateo del Mar Huave has *miyeed oniiügueran* ‘body.hair eye’ for ‘eyelash,’ and *miyeed opech oniiügueran* for ‘eyebrow.’

Other associations include: Mali *sachong angēt kēseng* is analyzable as ‘vision its those.particular.long.ones,’ while Oneida *oʔnekʰhtalaʔ* also means ‘strawberry top.’ The Norton Sound dialect of Central Yup’ik has *qeľemyaq* ~ *qeľemsaq*, containing a verb root meaning ‘to close eyes,’ while the Arabela term *maque-teja* is analyzable as ‘sleep/dream-INSTR.’ Rotuman *lekleki* also denotes a kind of tree.

### 113. The Eyelid

Representation: 71%

Motivated: 73.7%

Thereof Analyzable: 68.9%

Thereof Colexifying: 5.3%

Thereof by Contiguity: 29.4%

Thereof by Similarity: 20.8%

Recurrent associated meanings: eye, skin, peel/rind/shell, bark, leather/hide, eyelash, cover/lid, husk, lip, edge, surface, hair, scale, eyebrow, on, door/entrance, roof

Most frequently, motivated terms for ‘eyelid’ in the languages of the sample are analyzable and of the lexical type, consisting of elements meaning ‘eye’ and ‘skin,’ as in Yir Yoront *mel-pertn* ‘eye-skin.’ Such terms are (sometimes with additional meanings colexified with ‘skin’ to be discussed below) also attested in Efik, Hausa, Mbum, Ngambay, Dongolese Nubian, Berik, Kwoma, Sahu, Basque, Ket, Sora, Biloxi, Cheyenne, Kashaya, Kiowa, Lesser

Antillean Creole French, Pawnee, Pipil (Santo Domingo de Guzmán dialect), Yuki, San Lucas Quiaviní Zapotec, Copainalá Zoque, Aguaruna, Bororo, Carib, Cayapa, Hupda, Maxakalí, Miskito, Piro, Imbabura Quechua, Rama, Tsafiki, Wichí, Great Andamanese, Hani, Malagasy, Mandarin, White Hmong, Takia, Tetun, and Yay; Cubeo has a term derived from 'eye' by means of a classifier for coverings, shells, skin, etc., which however, also has a lexical counterpart of the same phonological form. Due to colexification with 'bark' in many languages (section 135), the Efik, Kwoma, Sahu, Basque, Biloxi, Pipil, Copainalá Zoque, Bororo, Cayapa, Cubeo, Maxakalí, Piro, and Tsafiki terms for 'eyelid' also betray an association with this meaning, and due to colexification with 'leather,' this meaning is associated in Dongolese Nubian, Sora, Lesser Antillean Creole French, Pipil, Copainalá Zoque, Cavineña, Miskito (which also colexifies 'feather'), Piro, Tsafiki, Wichí, and Wichí. However, for 'bark,' there are also terms where it is genuinely 'bark' rather than 'skin' which figures as a constituent of the term: Abzakh Adyghe has *na-p<sup>o</sup>e* 'eye-bark/shell,' and such terms, without colexification of 'skin,' are also found in San Mateo del Mar Huave and Sedang, and by direct colexification in Guaraní. Further, the Efik, Basque, Sahu, Itzaj, and Sedang terms also betray an association with 'husk,' and the Efik and Biloxi terms also with 'scale.' The association with 'rind,' 'peel,' and/or 'shell' is also present due to colexification with 'skin' and/or 'bark' in Efik, Basque, Itzaj, Lesser Antillean Creole French, Pawnee, Copainalá Zoque, Cayapa, Cubeo, Maxakalí, Piro, Tsafiki, and Hani, while in Bora, the term for 'eyelid,' *hállumiho*, is derived from *hállu* 'eye' by the classifier *-mi:ʔo* for hard shells and Hawaiian has *kuapo'i-maka*, with *maka* meaning 'eye' and *kuapo'i* "shell on back of crab or turtle" as well as 'kneepan,' among other meanings.

Hawaiian also has another term for the 'eyelid,' *ūpo'i maka* 'cover/lid eye,' and such terms (with 'cover' either being nominal or verbal) are also found among the languages of the sample in Dongolese Nubian, Sora, Welsh, Haida, Lesser Antillean Creole French, and Yana; the connection to a verb meaning 'to cover' is also etymologically detectable in Ineseño Chumash, and Piro colexifies 'lid' with 'surface' in general, for which compare Haida *xang ʔun* /*xang ʔunna*/ 'eye top/surface.' Chayahuita has *ya'pira yonsan* 'eye edge.' An analogous term is found in Yanomámi, and Hawaiian has *lihilihi*, reduplicated from *lihi* 'edge,' which colexifies also 'lace,' 'to crochet,' and denotes a variety of sweet yam. Kiliwa has *yuw=ha?=hiy* 'eye=mouth=edge/border,' with *yuw=ha?* meaning 'face' and *ha?=hiy* 'lip.' Similarly, Kosarek Yale has *heing bam* 'eye lip,' a term with such structure is also found in Wappo, while Cashinahua directly colexifies 'eyelid' with 'lip.' Ngambay has *dà-kèm* 'head/on-eye,' Nikvh *njaχ-t'xy* 'eye-on,' Buli *num-gbong* 'eye/platform-roof' and Bezhta *hāys ʔ'āq'e* 'eye.gen roof.' Curiously, One has *namna palla* 'eye body.hair,' and such a term is also found in Santiago Mexquititlan Otomí, where it colexifies 'eyelid' with 'eyelash,' as well as in Cavineña. Colexification of 'eyelash' and 'eyelid' is furthermore found in Swahili, Gurindji, Badaga, Kolyma Yukaghir (by the term *and'ə-šepil* 'eye door,' denoting 'upper eyelash' and 'upper eyelid' specifically; Dongolese Nubian also has a term where the second constituent next to 'eye' is 'door'), Upper Chehalis, Santiago Mexquititlan Otomí, Cavineña, Miskito, and Hawaiian (by the term *lihilihi* mentioned above).

Other complex terms for the ‘eyelid’ on the basis of ‘eye’ are Efik *mfut-ënyin* ‘shade eye,’ Kanuri *f̣rtà shîm-bè* ‘root eye-of,’ Kildin Saami *čall’m-rūmtas* ‘eye-brim,’ Blackfoot *ootokiáápinihpis*, perhaps containing *ootoki* ‘skull’ and *aapini* ‘eye,’ Cahuilla *pušyúmuve?*, literally ‘eye hat,’ Upper Chehalis *łúkw=ičn-s t mús* ‘above/top=ridge=?? INDEF.ART eye/eyeglasses’ for ‘upper eyelid’ specifically, Kiliwa *yuw=nat-u?=kw-waa* ‘eye=top/atop-OBL=WH-sit,’ Chayahuita *ya’pira sha’shatë* ‘eye corner.of.mouth,’ Embera *dau-í*, apparently analyzable as ‘eye-wing,’ Toba *l’ai lapo’té* ‘upper eyelid,’ containing *’ai* ‘eye’ and *apo* ‘poncho,’ *l’ai’té lqa* ‘lower eyelid,’ containing *l’ai’té* ‘eye’ and *lqa* ‘chin,’ and Fijian *daku-daku-ni-mata* ‘back-RED-POSS-eye’ for the ‘upper eyelid’ and *dreke-ni-mata* ‘hollow/cavity-POSS-eye’ for the ‘eye socket’ and ‘lower eyelid,’ Lenakel *nouanhal-nimr-* ‘egg-eye-,’ and Manange *2mi-2p<sup>hi</sup>* ‘eye-up.’ Finally, Tetun has *matan-kukun* ‘eye-nail/claw,’ and due to colexification with ‘skin,’ the association with ‘fingernail’ is also found in Biloxi. Semianalyzable terms where the identifiable constituent is ‘eye’ are found in Bakueri, Yoruba, Toaripi, Waris, Sora (for the ‘lower eyelid’ specifically), Carrier, Upper Chehalis, Highland Chontal, Comanche, Santiago Mexquititlan Otomí, Xicotepec de Juárez Totonac, Wappo, Cavineña, Hani, and Rotuman.

Furthermore, the Oneida term *okahkwilo?ókta?* likely revolves around the verb root *-kahkwilo?ok-* ‘to blink,’ and similarly, Khalkha *anisxa* may be related to the verb *aniski-* ‘to blink eyes repeatedly.’

Other associations include: Muna *bhangkoa* also means ‘eye socket’ and, verbally, “to roll up something broad but thin.” *Zobki*, a Khalkha term for ‘eyelid,’ also means ‘corner of the eye’ (compare *zobkis* ‘corner’). Ineseño Chumash *’ixma’y* also means “to have a film or cloud in the eye, cataract” and Guaraní colexifies ‘eyelid’ with ‘petal.’

#### 114. *The Finger*

Representation: 96%

Motivated: 59.6%

Thereof Analyzable: 22.1%

Thereof Colexifying: 37.5%

Thereof by Contiguity: 20.5%

Thereof by Similarity: 34.7%

Recurrent associated meanings: hand, toe, arm, claw/paw/forefoot, child/son, branch out, little, fingernail, branch, fruit, fork, hoof, head, neck

‘Finger’ and ‘hand’ are often not distinguished lexically (Brown 2005a), and this is also true of many sampled languages, namely Ngambay, Burarra, Kyaka, Ngaanyatjarra, Nunggubuyu, Tasmanian (Western, Middle-Eastern and Southeastern), Yir Yoront, Abzakh Adyghe, Cahuilla, Upper Chehalis, Cheyenne, Kashaya, Kiliwa, Nez Perce (where the narrow interpretation ‘finger’ is restricted to compounds), Oneida, Wintu, Yana, Arabela, Cavineña, Chayahuita, Jarawara, Great Andamanese, and Hawaiian. In a subset of these languages, the same term is also used for the ‘arm’ as a whole. This is the case in Ngambay, Kyaka, Ngaanyatjarra, Cahuilla, Kiliwa, and Hawaiian, where the relevant term also means ‘sleeve’ and ‘five.’ However, it is very often also the case that there are complex terms for ‘finger,’ often metaphor-driven, where one of the constituents is ‘hand,’ which is sometimes colexified with ‘arm’ generally (Brown 2005b). None of these metaphor-driven pat-

terns is overwhelmingly frequent, but they do recur on a global scale with some strength. The most common meaning of the second constituent in complex terms is ‘child,’ as in Katcha *bibala ma nizo* ‘child GEN hand.’ Such terms are also found in Koyraboro Senni, Mbum, Gurindji, Chickasaw, Itzaj, Pipil, and Maxakalí (in Chickasaw and Pipil with ‘son’ specifically rather than ‘child’ generally). Perhaps relatedly, in Yana, *tid?yau(na)* is also a kinship term for ‘man’s younger brother’ and ‘woman’s younger sister.’ A word meaning ‘small’ or ‘little’ also figures in some languages, for instance in Miskito *mita sirpi* ‘hand little.’ Similar terms are found in Carib, Piro, and Samoan. In Mali, the word for ‘finger’ is derived from that for ‘hand’ by means of a masculine noun class prefix, which is associated with smallness (Stebbins 2005), and Yir Yoront, which colexifies ‘finger’ and ‘hand,’ has the optional complex term *yor-mart* ‘hand/finger-little/tiny’ to single out the ‘finger’-reading. In the Piro term, the second constituent conveying ‘smallness’ has also a reading ‘fruit’ (the association between diminutive semantics and ‘fruit’ is also reported by Matisoff 1992). Similarly, Lenakel has *noua-nelm-* ‘fruit-hand/arm-,’ and Tetun *liman-fuan* ‘arm/hand-fruit/heart.’ In the Carib term, the constituent *sikili* is glossed more precisely as ‘little end,’ and similarly, Kapingamarangi has *madaalima* /*mada-lima*/ ‘end-hand/arm.’ Hawaiian *manamana*, colexifying ‘finger’ with ‘claws, rays, forks’ inter alia, is reduplicated from *mana* ‘branch,’ Malagasy has *rantsantànana*, analyzable as /*ràntsana-tànana*/ ‘branch-hand,’ and ‘branch’ and ‘finger’ are colexified directly in Aguaruna. Relatedly, Nuuchahnulth has *čačəłəqniukum* /čə-čəłəq-nuk<sup>w</sup>-im/ ‘RED-branch.out-in.or.at.the.hand,’ and similar terms highlighting that the fingers protrude from the hand are featured in Kaluli and Yay (*dagi gasa* ‘hand/paw things.that.separate’ and *jiaj<sup>5</sup> fu<sup>4</sup>* ‘to.come.apart hand’ respectively) alongside a reduplicated term with the meaning ‘branch, branch out’ in Hawaiian. Similarly, Tehuelche and Hawaiian colexify ‘finger’ with ‘fork.’ Pawnee has *iks-kiic-u?* ‘hand-neck-NOM,’ and Bororo *era-ko* ‘hand-neck.’ Other complex terms where one of the constituents is ‘hand’ and/or ‘arm’ are Noni *kpwān ebo* ‘to.mark hand,’ Baruya *a’jawinya* /*ata-jawinya*/ ‘hand-fish’ (also denoting the ‘thumb’ specifically; note that in other languages such as Khoekhoe, ‘thumb’ is explicitly not included in the extensional range of the ‘finger’-word), Muna has *wuna-no lima* ‘flower-poss hand/arm,’ Sko *nò-kang-kang* ‘arm/hand-tusk/eat-RED,’ Abzakh Adyghe *?e-pe* ‘hand-tip’ (compare possible connections to this meaning in Ancient Greek as well as cognates between Baltic and Germanic suggestive of the association, Buck 1949: 240), Bora *méhójsiwa* /*mé-hójsiŋ-gwa*/ ‘1PL-hand-SCM.2d.straight,’ Cubeo *āmu-yo* ‘hand/arm-CLASS.LARGE.CYLINDRICAL.SLIM.AND ACUTE.OBJECT,’ Cayapa *tyamishu* /*tyaapa-mishu*/ ‘hand-head,’ a term with identical structure in Tsafiki, Kaingang *nīgé féj* ‘hand leaf’ and *nīgé juféj* ‘hand long.things,’ and Bwe Karen *-cu-yò* ‘-arm/hand-point.at’ (note also that Rotokas *piiroo* also means “point towards something with finger or object;” this association is probably also at the heart of Latin *digitus*, Buck 1949: 240).<sup>9</sup> Furthermore, Yanomámi *imi-hena* is analyzable as ‘carry.in.hand-leaf.’ Semianalyzable terms are featured in Katcha, Dadibi, One, Sentani, Abzakh Adyghe, Biloxi, Carrier, San Mateo del Mar Huave, and Wappo.

<sup>9</sup> To this list of metaphorical denominations one can add, from outside the sample, Dene Sųliné’s *denelatthalé* ‘person’s hand-awl’ for ‘finger’ and *denelachédh* ‘person’s hand-duck’ for thumb (Rice to appear).

Furthermore, Ngaanyatjarra, Yir Yoront, Basque, Upper Chehalis, Cheyenne, Lengua, Rama, Hawaiian, and Rotuman colexify ‘finger’ with the homologous structures in animals, namely ‘claw,’ ‘paw,’ or ‘forefoot.’ Similarly, in the Nunivak Island dialect of Central Yup’ik, *ipik* ~ *ipi* may also be used with reference to the “limb of quadruped or insect.” Kwoma colexifies ‘finger’ with ‘person’ (as well as, by extension, ‘twenty’ and other meanings), and the Hooper Bay and Chevik dialect of Central Yup’ik has *cugaraq* ~ *cuarraq* ~ *yuarraq* containing *cuk* ~ *yuk* ‘person.’ Basque *behatz*, containing *hatz* ‘finger,’ also means ‘hoof,’ ‘vestige, trace, mark’ and other things, (‘hoof’ is also colexified in Lengua), while the Buli and San Lucas Quiaviní Zapotec terms are also traditional units of measurement. Hausa, Dadibi, and Basque colexify ‘finger’ with ‘fingernail’ (Hausa also with “[d]ermatoid cyst under upper eyelid of horse”).

In addition, when simplex terms for ‘finger’ do exist (sometimes exhibiting other patterns of colexification as well), they can often at the same time be used with reference to the ‘toe’ (also in some Indo-European languages, Buck 1949: 240-241), with disambiguating complex terms in some languages. This is true of Buli, Efik, Rendille, Swahili, Yoruba, Kwoma, Lavukaleve, Rotokas, Northeastern Tasmanian, Toaripi, Sahu, Waris, Badaga, Bezhta, Japanese, Ket, Khalkha, Welsh, Blackfoot, Cheyenne, Ineseño Chumash, Haida, Tuscarora, Central Yup’ik, San Lucas Quiaviní Zapotec, Copainalá Zoque, Aymara, Cayapa, Huambisa, Lengua, Macaguán, Fijian, Hawaiian, Lenakel, Rotuman, Sedang, Takia, and White Hmong. An interesting generalization emerges, namely, that terms betraying an association between ‘finger’ and ‘hand’ and those colexifying ‘finger’ and ‘toe’ are mutually exclusive in the languages of the sample. That is, there is no language in the sample in which all three meanings are colexified (although one language may feature several terms for ‘finger,’ one of which features one association and another one the other). This thus appears to be a strongly dispreferred pattern cross-linguistically.

Other associations include: Berik *bola* colexifies ‘forefinger’ with ‘ball,’ Kyaka *kingi* also may refer to “name, title, rank, reputation,” and Ngaanyatjarra *mara* also means ‘five.’ Chukchi *rəlyəlyən* contains *yalən* ‘skin,’ Welsh *bys* colexifies ‘hand of clock,’ and Cheyenne *mo’ěsko* ‘ring.’ Tuscarora *uhsú?kweh* may also be used to refer to a ‘rake,’ a ‘thimble’ or a ‘herring tooth.’ The bound term *wa-* in Yana is also a general prefix for long objects, while Jarawara *yehe/yehe* may also refer to the ‘self.’ Rama *isúluk up* contains *up* ‘eye, something round’ and presumably *uk* meaning ‘skin, coat, rind’ inter alia. *Ehethaki*, a Yanomámi term for ‘finger,’ consists of *ehetha* ‘wrist’ and the quantal classifier *ki* (for which see § 4.4.1). Rotuman *kākā’e* also means ‘wall of house,’ and Lesser Antillean Creole French *dwet* also means ‘upright, straight, direct’ and, by extension, ‘righteous.’

### 115. The Fingernail

Representation: 95%

Motivated: 69.2%

Thereof Analyzable: 12.8%

Thereof Colexifying: 56.1%

Thereof by Contiguity: 1.3%

Thereof by Similarity: 46.4%

Recurrent associated meanings: claw/talon, toenail, hoof, finger, hand, shell, arm, skin, paw, bark, horn

'Fingernail' is very frequently colexified with 'claw' and/or 'talon' (see Buck 1949: 245 on Polish). This is the case in as many as 67 sampled languages, namely Bakueri, Buli, Efik, Hausa, Khoekhoe, Noni, Dongolese Nubian, Rendille, Swahili, Anggor, Buin, Burarra, Gurindji, Mali, Ngaanyatjarra, Toaripi, Waris, Yir Yoront, Badaga, Basque, Bezhta, Greek, Japanese, Ket, Khalkha, Laz, Kildin Saami, Welsh, Kolyma Yukaghir, Acoma, Upper Chehalis, Cheyenne, Highland Chontal, Ineseño Chumash, Haida, Itzaj, Lake Miwok, Lakota, Lesser Antillean Creole French, Kiowa, Oneida, Pipil, Xicotepec de Juárez Totonac, Tuscarora, Wintu, Yaqui, Yuki, Central Yup'ik, San Lucas Quiaviní Zapotec, Copainalá Zoque, Aguaruna, Aymara, Carib, Cashinahua, Cayapa, Cubeo, Guaraní, Huambisa, Miskito, Piro, Ancash Quechua, Wayampi, Fijian, Hawaiian, Malagasy, Rotuman, and Tetun (in Rotuman, by a semianalyzable term containing 'hand'; there are other such terms in Upper Chehalis and Hani). Complex terms betraying this association are found in Wappo *meʔčúš*, containing *meʔ* 'hand, by hand' and *čúš*, also occurring in *pheʔčúš* 'claw, hoof, toenail,' and in Great Andamanese *ôngkôro-bôdoh* 'finger/hand-claw.' There is a semianalyzable term in Baruya. In a subset of these languages, Efik, Dongolese Nubian, Toaripi, Greek, Kildin Saami, Biloxi, Cheyenne, San Lucas Quiaviní Zapotec, Aymara, Arabela, Guaraní, Miskito, Ancash Quechua, Fijian, Hawaiian, Rotuman and Tetun, 'hoof' is in addition colexified; this is also the case in Muna and Basque. Furthermore, Buli, Guaraní, and Miskito colexify 'paw' and 'fingernail' in addition.

Biloxi and Rama colexify 'fingernail' with 'bark,' and a complex term is found in Jarawara (*yehe atari* 'finger/hand skin/bark/peel'). Biloxi also colexifies 'fingernail' with 'horn' (also 'scales of fish'), and Abzakh Adyghe has *ʔebžane* /ʔe-bž(e)-ne/ 'hand-horn-part/unit' (this association is also found Avestan and Ancient Greek, Buck 1949: 245). Unsurprisingly, the complex terms so far mentioned are not the only ones where one of the constituents is 'hand' or 'finger' (or 'arm'). One has *meu alo* 'arm egg.shell,' Toaripi *mai haro* 'hand/arm head/shell,' Sko *nòbi* /nò-bí/ 'arm/hand-shell,' Hupda *cob b'ók* 'finger-skin/shell,' Bislama *sel fingga* 'shell finger,' Samoan *atagi lima* 'shell arm,' and Yir Yoront colexifies 'fingernail' with various kinds of shell directly (note also that Hausa *k'umba* also denotes the 'shell of a mussel' inter alia). For the association with 'skin,' note also that Biloxi and Rama colexify 'fingernail' and 'skin,' as well as the complex terms *tofweftw'oj*, containing *fweftw* 'finger' and *t'oj* 'skin, hide' in Wichí and *imi-si* 'carry.in.hand-cover/skin' in Yanomámi. There is a semianalyzable term involving a constituent 'skin' in Chukchi, and a semianalyzable one with a constituent meaning 'skin, bark, shell' in Guaraní. Other complex terms are Kwoma *tapa bi* 'arm/hand point,' Blackfoot *awó'taanookitsis* /awó'taan-mookitsis/ 'shield-finger/toe,' Tsafiki *tehué* /tede-hué/ 'hand-fast,' Manange *1ja-2fij* 'hand-wood,' and Mandarin *zhi3-jia1* 'finger-armor.' Hausa, Dadibi, and Basque directly colexify 'finger' and 'fingernail,' and semianalyzable terms involving either 'finger,' 'hand,' and/or 'arm' are furthermore attested in Mbum, Dadibi, Kwoma, Carrier, Chickasaw, Copainalá Zoque, Cayapa, Maxakalí, and Rotuman.

Finally, as is the case with 'finger' and 'toe' (compare section 114), 'fingernail' and 'toenail' are often colexified, at times with the possibility of disambiguating complex terms. Terms explicitly glossed as also capable of referring to the 'toenail' are found in Buli, Efik, Khoekhoe, Rendille, Anggor, Buin, Gurindji, Kaluli, Kyaka, Lavukaleve,

Ngaanyatjarra, Nunggubuyu, Rotokas, Tasmanian (Northeastern, Middle-Eastern, and Southeastern), Sahu, Waris, Basque, Chukchi, Ket, Kildin Saami, Sora, Biloxi, Blackfoot, Upper Chehalis, Cheyenne, Highland Chontal, Itzaj, Oneida, Wintu, Central Yup'ik, Cayapa, Embera, Piro, Bislama, Fijian, Hawaiian, Lenakel, Rotuman, and Tetun.

Other associations include: Nunggubuyu *yara* also means 'to smell, detect, sense something,' while Muna *konisi* also denotes a kind of crab. Ngaanyatjarra *miltji* also denotes the "spines of thorny devil" and means 'rake.' The Kosarek Yale term *aluk* also is used as the name of a variety of banana, and Yir Yoront *yorwel* also means 'spoon.' Greek *níchi* is also used figuratively with the meaning 'clutches,' and Ket colexifies 'fingernail' also with 'needle, pin.' Cahuilla *sálul* is presumably related to the verb *-sáluk* 'to scratch,' the Nuuchahnulth term *čač̓a* contains the verb *čač̓-*, meaning 'to get split lengthwise,' while Wintu *kah* also means 'wind, blow.' Wayampi *pāpē* contains *āpē* 'to bend' and also denotes 'vaned feathers,' Fijian *kuku* is also the name for a kind of mussel, while Bwe Karen *θāmi* is also the name of a shell fish species. The Kapingamarangi term *madaniha* consists of *niha* 'tooth' and the augmentative prefix *mada-*, and Tetun *kukun* colexifies 'fingernail' with 'pincer of crab.'

#### 116. *The Guts*

Representation: 95%

Motivated: 47.7%

Thereof Analyzable: 13.0%

Thereof Colexifying: 34.9%

Thereof by Contiguity: 3.9%

Thereof by Similarity: 5.9%

Recurrent associated meanings: innards/bowels/entrails, belly/stomach, faeces, inside(s), rope/string/line, sausage, heart, womb, catgut, child/son, inner tube of tire, end

Many sampled languages have terms referring to the 'innards' or 'entrails' or 'bowel(s)' generally, including the 'guts' or 'small/large intestine.' These are Buli, Efik, Khoekhoe, Mbum, Ngambay, Dongolese Nubian, Yoruba, Buin (the term includes also the 'bladder' in its semantic range), Kyaka, Lavukaleve, Muna, Nunggubuyu, Sentani, Kosarek Yale, Yir Yoront, Badaga (where the relevant term is also capable of referring to the 'liver' specifically), Greek, Khalkha, Welsh, Ineseño Chumash, Haida, Lakhota, Lesser Antillean Creole French, Nez Perce, Pawnee, Tuscarora, Central Yup'ik, Abipón, Aguaruna, Arabela, Cashinahua, Embera, Toba, Fijian, Bislama, Bwe Karen, Hawaiian, Rotuman, and Samoan; in the latter two languages, the term is also a jocular designation for a 'child' or 'son' specifically. 'Guts' and 'inside(s)' are colexified in Kaluli, Kwoma, Abzakh Adyghe, Yuki, Jarawara, Lengua, Lenakel, and Sedang, while Embera colexifies it with 'innermost,' and Japanese has *nai-zō* 'inner-organ.'

As for other associations with body-parts, Badaga, Abipón, and Hawaiian colexify 'guts' with 'heart' (among other meanings in some languages, see section 117); 'heart' is in quotation marks in the Badaga source, suggesting a figurative rather than literal usage, and the same may be true of Hawaiian. Sahu, Badaga, and Sedang colexify 'guts' with 'womb' (the latter language also with "inner edge of bamboo or of kōmea square strip of rattan").



In some sampled languages, there is a metaphorical transfer realized by analyzable terms of the lexical type from 'line' or 'rope' to 'guts.' For instance, Hupda has *tok tít* 'stomach string/vine,' and in Toaripi, 'rope' and 'intestines' are colexified. Also attested in this class of terms is to have 'faeces' rather than 'belly' or 'stomach' as the second element, as in Cayapa *pe-shilli* 'excrement-line.' Such terms are also attested in San Mateo del Mar Huave and Tsafiki. Nuuchahnulth *čiyup* contains *čis* 'in line, rope,' and a semianalyzable term involving a constituent meaning 'thread, string' is found in Highland Chontal.

Khoekhoe, Mbum, Kwoma, Ngaanyatjarra, Southeastern and Western Tasmanian, Badaga, Basque, Khalkha, Cheyenne, Kiliwa, Lake Miwok, Guaraní, and Miskito directly colexify '(part of the) guts' with 'stomach' and/or 'belly' (Badaga also with "something that came from the stomach"). Burarra more specifically colexifies 'stomach' with 'part of intestines.' Katcha has *em(a) aidhi* /*ema aidhi*/ 'object stomach,' Meyah *otkonú efesá* 'stomach child,' Carib *uwempo sikilipo* containing *uwempo* 'belly' and *sikili* 'little parts,' and Toba *lailshic*, presumably analyzable as /*l-ael-shi-c*/ '3SG.POSS-stomach-CLASS.PLACE.OF.PROVENIENCE-MASC' and *l-ael l-ec* '3SG.POSS-stomach 3SG.POSS-content.' Furthermore, in Swahili, the term for 'guts' consists of that for 'stomach' and a noun class prefix, in Embera, the relevant meanings are expressed by the same term with alternating genders, and semianalyzable terms where the identifiable constituent means 'belly' and/or 'stomach' exist in Sko and Bora. Ngambay, Baruya, Tasmanian (all varieties), Haida and Jarawara colexify 'guts' with 'faeces' directly (Baruya also with 'flatulence'), Kosarek Yale has *disidik*, where *di* is 'faeces' and *sidik* can mean 'rest, leftover,' 'genuine, true,' as well as 'to straighten, sprout, grow quickly,' Bororo has *pe-guru* 'excrement-liquid,' Cubeo *cúra-me* 'faeces-CLASS.LIKE.THREAD,' Piro *hitška-pi* 'excrement-rod,' and in Yanomámi, *shi-ki ~ shi-pě* consists of *shi* 'faeces' and a quantal classifier (see § 4.4.1.1.). Semianalyzable terms where the identifiable constituent means 'faeces' are attested in Berik, Kashaya, Pipil, Sáliba, and Lenakel, and ones where it means 'to defecate' in Abzakh Adyghe and Chayahuita.

Abzakh Adyghe also has the term *č'et'ay* /*č'e-t'ay(e)*/ 'end-loop,' and Kolyma Yukaghir *jömgid-ičə* 'turn-end.' Presumably by provenience contiguity, Bezhta, Nez Perce, Tuscarora, Manange, and Rotuman colexify 'guts' with 'sausage' (the association was present in Latin by formation of a diminutive, Buck 1949: 1086), and, by metaphor, Buli and Hausa colexify 'guts' with 'inner tube of tire' (Hausa also with 'works,' e.g. as of a clock), and Basque and Lesser Antillean Creole French with 'catgut' (Basque also with a 'medicinal probe').

Other associations include: Khoekhoe *!nāb* also means 'interior' and 'diarrhoea.' Yoruba *orièkọ* contains *èkọ* 'head,' Dadibi *haliga* also means 'umbilical cord,' and Kwoma *sugu* is also the name of a 'digging stick.' Muna *ghule* also means 'snake, worm' and *ngallungallu* also denotes a "rattan nose-ring for cows or buffaloes." Ngaanyatjarra *tjuni* colexifies 'guts' and 'stomach' with 'front of person,' the "hollow of anything concave," the 'round part of fruit,' and is conceived of as the seat of emotions (which is also the case in Hawaiian), while Nunggubuyu *-lhan<sup>9</sup>aj* also denotes a "minnow, small fish (used as bait)." Another Nunggubuyu term, *-n<sup>9</sup>u-*, also means 'to eat,' while Waris *ungul* 'large intestine' appears to contain *ung* 'anus.' Basque *golko* also means 'bosom, breast' as well as 'bay,

gulf,' Japanese *chō* also means 'butterfly' (it is in both readings borrowed from Chinese, but goes back to different etyma), and Khalkha *gedesy(n)* is also the name of the "loop for o/u in Khalkha script;" this term may be derived from *gede* 'nape of neck, occiput.' Sora *on'loida:n ~ onlaida:n* contains *l'oj-* 'testicle.' Welsh *perfedd* also means 'middle,' and *pot* also 'navel.' Central Yup'ik *cakunglluut* contains the postbase (see § 4.4.2) *-nglluk* 'not too good,' and *imanaq imaq* 'contents, bullet, pus, ocean' and perhaps the postbase *-naq* 'one like.' Arabela *mana-ca* is analyzable as 'mud-CLASS.FRUIT,' Jarawara *tori/toro* also means 'streambed,' Wayampi *ikε* also 'side' and 'older sister,' Fijian *wāwā* also 'tired out,' and Bwe Karen *-bwì* also for beverages "to be potent, strong."

### 117. *The Heart*

Representation: 97%

Motivated: 48.0%

Thereof Analyzable: 11.5%

Thereof Colexifying: 37.3%

Thereof by Contiguity: 4.7%

Thereof by Similarity: 22.6%

Recurrent associated meanings: feel/think, center/middle/core, soul/spirit, belly/stomach, inside/interior, heart in cards, liver, courage/boldness, life/alive, breast, seed/pit, pith, lungs, breath, want/desire, kidney, love, conscience, fruit, guts, womb, wind

Terms for 'heart' are sometimes lexically connected to other internal organs of the body on the one hand, and are, on the other hand, associated with emotional states and cognitive abilities more generally in many languages.

By colexification, there are three languages, Ngambay, Guaraní, and Yanomámi, where 'heart' can also refer to the 'kidney' (and in Yanomámi to a spherical object in general), and in four, Buin, Burarra (by the analyzable term *mun-molma* 'CLASS.DOMESTIC-warmth'), Kwoma, and Yay, 'heart' and 'lungs' are colexified respectively (and there are other languages in which the latter meanings are expressed by morphologically complex terms on the basis of 'heart,' see section 122). Colexification with 'liver' is a little more frequent, occurring in Ngambay, Kwoma, Sahu, and San Lucas Quiaviní Zapotec, while in Yay, *tuay*<sup>2</sup> denotes the internal organs 'heart' and 'lungs' and sometimes includes the 'liver' (note also that terms similar in shape to Nunggubuyu *andhīri* 'heart' are found in neighboring languages with the meaning 'liver'). Moreover, in Khoekhoe there is a very general term for 'offals,' including 'lungs, liver, kidney' and 'heart' that can also refer to the 'belly' or 'stomach.'

In addition, there is one language, Mbum, where 'heart' is *làù wârké* 'liver male.' More frequent is an association with the 'stomach' cross-linguistically. Ngambay, Abzakh Adyghe, Badaga, Laz, San Mateo del Mar Huave, Itzaj (by the term *pusik'al*, analyzable as /puus-ik'-al/ 'dusting-wind-COLL;' for this, note also Cubeo *ũme-du* 'wind-CLASS.ROUNDISH.THREEDIMENSIONAL.OBJECT' and that the association is present in Tariana according to Aikhenvald 2003: 128), Santiago Mexquititlan Otomí, Yuki, Arabela, and Guaraní colexify the meanings, in which case the respective terms appear to be more vague in their reference, referring to the internal part of the trunk more generally. Similarly, Piro

colexifies ‘pit of the stomach’ more specifically, and furthermore, a semianalyzable term featuring a constituent meaning ‘stomach, belly’ as well as ‘front of body’ exists in Yir Yoront. Badaga, Abipón, and Hawaiian colexify ‘guts’ with ‘heart’ (Badaga also with ‘dysentery,’ ‘womb,’ and other meanings; note though that ‘heart’ is in quotation marks in the Badaga source, which suggests a figurative rather than literal usage, and the same may be true of Hawaiian), Katcha, Highland Chontal, Kaingang, and Great Andamanese with ‘breast,’ and Badaga and Welsh with ‘womb,’ although this reading is obsolete in Welsh. Moreover, Kwoma colexifies ‘heart’ with the “upper half of torso,” and Sahu with ‘chest.’ For the associations with ‘wind’ in Itzaj and Cubeo just mentioned, note also that in some languages there is a connection with ‘breath’ or ‘breathe.’ Kolyma Yukaghir, Wintu, and Rotuman colexify these meanings (Wintu also ‘to rest, be restored, revive’), and Rotokas has *vovou isi* ‘breathe/want CLASS.ROUND.OBJECT.’

On the other hand, the heart is often conceived of as the seat of emotions or the soul. Yoruba, Lavukaleve, Sahu, Badaga, Kolyma Yukaghir, Yaqui, San Lucas Quiaviní Zapotec, Guaraní, Hupda, Rama, Tsafiki, and Bwe Karen colexify ‘heart’ with ‘soul’ and/or ‘spirit’ (Yaqui also with ‘ghost’), and in as many as 33 sampled languages, ‘heart’ is lexically associated with cognitive abilities, such as ‘feeling,’ ‘thinking,’ or is considered the seat of emotions more generally. This is the case by colexification in Noni, Buin, Burarra, Kyaka, Lavukaleve, Muna, Sahu (where the relevant term also means “to tell a story”), Waris, Badaga, Khalkha, Haida, Nuuchahnulth (where the relevant term also means ‘brain’ and ‘spinal cord’), Xicotepec de Juárez Totonac, Guaraní, Tehuelche, Tsafiki, Hani, Bwe Karen, Hawaiian, Lenakel, Mandarin, Rotuman, and Sedang, and by complex terms in Buli (*sukiri* ~ *sikiri* /sui-kiri/ ‘mind-root/reason’), Upper Chehalis (*s-qwálm* ‘CONTINUATIVE-think’), and Abipón (*n-eo-et-Ra-nat-Ra* ‘POSS.INDEF/3SG-feel-COGNITIVE.ACTIVITY-ABSTR-??-ABSTR’). There is a semianalyzable term where the identifiable constituent also means ‘middle’ in Buin. In four languages of the Old World, Efik, Badaga, Khalkha, Welsh, and also in Lesser Antillean Creole French, associations with ‘courage,’ ‘boldness’ and other meanings are found (see Buck 1949: 251 for Indo-European), and in three languages, Waris, Bora, and Ancash Quechua, the heart is lexically associated with ‘love.’ In Basque, it is also a term of endearment (glossed as ‘darling’). Alongside Rotokas, Khoekhoe has *tgao-b* ~ *tgao-s* ‘want/desire-3SG.MASC’ ~ ‘want/desire-3SG.FEM,’ and such associations are found by colexification in Muna, Bwe Karen, and Rotuman, and a semianalyzable term with an element meaning ‘to breathe’ as well as ‘to want, desire’ exists in Rotokas. Note also that in Kiowa, where ‘heart’ is *t’eqin*, there is also a verbal prefix *t’eqin-* “referring to desire.” Moreover, Yoruba, Badaga, and Guaraní colexify ‘heart’ also with ‘conscience.’ Other associations with cognitive and emotional states are: Buli *kpa-ziim* ‘occiput-blood’ means ‘heartbeat’ and ‘anxiety, fear’ and in restricted contexts also ‘heart,’ Hausa *zuciya* also means “to get into a temper,” “quick temper,” whereas Ngambay *wùr*, in contrast, also means ‘patience.’

The heart is also at times (in Sora, Nuuchahnulth, Tuscarora, Central Yup’ik, where the relevant term however is especially used for the heart of fish, and Bwe Karen) lexically connected with words for ‘life’ (in Sora also with ‘intimacy’ and ‘friendship’). This is particularly frequent in North America, where the connection is typically realized by morphologically complex terms (Nuuchahnulth, Tuscarora, and Central Yup’ik). For in-

stance, a dialectal Central Yup'ik word for the 'heart' is *unguvan* /*unguva-(u)n*/ 'be.alive-device.for.' The association occurs by colexification in Sora and Bwe Karen.

The 'heart' is also sometimes (namely in Khoekhoe, Yir Yoront, Basque, Kiliwa, Pawnee, and Bislama) extended to mean 'heart in cards,' likely under European influence. Frequently, the heart is also extended to mean 'center, middle' or 'core' (see Buck 1949: 251 on Indo-European and Heine and Kuteva 2002: 171 for a perspective from grammaticalization). This is the case in Abzakh Adyghe, Basque, Welsh (where the same term also means 'navel'), Kolyma Yukaghir, Ineseño Chumash, Itzaj, Lesser Antillean Creole French, Tuscarora, San Lucas Quiaviní Zapotec, Arabela, Bora, Lengua, Ancash Quechua, Mandarin, and Samoan, while in Buin, there is a semianalyzable term for 'heart' where the identifiable constituent is 'middle' (or 'to know, think of, reflect on'). Similarly, 'heart' is colexified with 'inside' or 'interior' in Ngambay, Northeastern Tasmanian, Basque, Maxakalí, Ancash Quechua, Malagasy and Rotuman, while the San Mateo del Mar Huave term mentioned above is analyzable as *omeaats-aran* 'inside-INAL.POSS.' Shape-based similarity, as well as similarity in the position of the two with respect to the larger structure they occur in is also responsible for associations between the heart and the meanings 'seed' or 'pit' in Sentani, Kosarek Yale, Nez Perce, and Samoan; note also that Hupda *hā'wig* is probably from *\*haŋ-wig* 'breath-seed.' However, 'fruit' itself is also colexified with 'heart' in Muna and Tetun (in Muna also with 'banyan tree'), and in Kosarek Yale in fact all three meanings can be expressed by a single term (as well as 'flower-stalk,' 'egg,' and figuratively 'child'). Further, Dadibi has a complex term for 'heart' involving *du* 'hill, mount, fruit.' In Ineseño Chumash, Nez Perce, Nuuchahnulth and Fijian, respective terms also can be used to refer to the 'pith' or 'center' of trees (the relevant Fijian term also denotes the breadfruit tree and its fruit).

Other associations include: the relevant Dongolese Nubian term is also used as a reflexive pronoun. Lavukaleve *vuvul* is also used to refer to a 'young coconut,' and Kosarek Yale *wangka* also in dancing songs to the "seed of sweet potato" specifically as well as, without restrictions as to register, to the "cocoon of a spider species." Abzakh Adyghe *gʷə* is extended to 'surrounding,' and, from there on, to 'surface, territory.' Kolyma Yukaghir *šubed'ə* colexifies "middle part of a fish trap, middle part of a boot" (this term is connected to a root meaning 'to run'). Cheyenne *hěsta* is also the name of the 'heart,' a constellation of stars. Haida obsoletely colexifies 'heart' with 'throat,' while the relevant Wappo term contains an element meaning 'stone.' Aguaruna *anentái* is derived from *ānen* ~ *āneg* 'love song' by means of the instrument nominalizer *-tai*, and Bororo *bapo* ~ *wapo* also denotes a 'rattle.' The Chayahuita term *nino-pi* is analyzable as 'beehive-CLASS.FRUIT.BODY.OR.BACK,' while Jarawara *makawari* appears to contain *maka* 'snake, jungle monster' and *wari* 'tree.' Ancash Quechua *puywan* ~ *puywaq* is derived from *puyway* 'for the heart to beat.' Imbabura Quechua *tulpa rumi* is analyzable as 'hearth stone,' whereas the Sáliba term *omaĩdi* also means 'animal.' Wayampi *tulu-ākā* appears to be analyzable as 'big-head,' Great Andamanese *ôtkûktâbana* contains *bana* 'globular,' Bislama colexifies 'hat, helmet' (due to phonological collapse of English *heart* and *hat*) and Hawaiian *pu'u-wai* is literally 'protuberance-water' and colexifies "a heart-shaped locket as of gold ... or silver," "small suckers on a taro plant," and has other figurative usages. Hani *neesiq* colexifies 'heart' with 'soy bean'

(*siq* is a classifier for round objects), Kapingamarangi *manawa* also carries the meanings ‘predisposition’ as well as to “come up from the surface (from diving),” Mandarin *xin1* (reflecting early middle Chinese *sim* and *sin* respectively) also means ‘new,’ and Sedang *ihiam* also denotes the “inside works of a machine” *inter alia*, and *nuih* also the “pole in roof of house.”

#### 118. *The Jaw*

Representation in Database: 79%

Motivated: 57.6%

Thereof Analyzable: 20.4%

Thereof Colexifying: 37.3%

Thereof by Contiguity: 50.1%

Thereof by Similarity: 4.1%

Recurrent associated meanings: chin, cheek, mouth, tooth, bone/skeleton, palate, gums

‘Jaw’ (sometimes ‘lower jaw’ specifically) is colexified with ‘chin’ in many languages in the sample (see also Buck 1949: 220 for relations between ‘jaw,’ ‘chin,’ and ‘cheek’ as well as, more rarely, ‘beard’ in Indo-European), namely in Yoruba, Baruya (by the analyzable term *maanaginya* /*maanga-yaginya*/ ‘mouth-bone’), Kaluli, Kwoma (by the term *teekibi*, perhaps containing *teeki* ‘to string, tense’ and *bi* ‘point’), Kyaka, Lavukaleve, Mali, One, Southeastern Tasmanian, Toaripi, Kosarek Yale (where the relevant term colexifies “wattle of an agama”), Abzakh Adyghe (by the analyzable term *žepqʷ* /*že-pqʷ(ə)*/ ‘mouth-skeleton’), Greek, Japanese, Ket (by a semianalyzable term containing it ‘tooth;’ such a term is also found in Embera), Khalkha, Nivkh, Welsh, Kolyma Yukaghir, Blackfoot, Upper Chehalis, Cheyenne, Chickasaw, Itzaj, Kiliwa (by a term containing *yu=wha?* ‘face,’ which is literally ‘eye=mouth’), Nez Perce, Nuuchahnulth (by the analyzable term *hiiniiłaksut* /*hiiniiłaxa-łaksut*/ ‘lower/below-at.the.mouth/at.the.lips,’ colexifying also ‘lower lip’), Pipil (by the analyzable term *-te:n-tsi:ka-w* ‘-mouth-ant-POSS’), Arabela, Aymara, Cashinahua, Embera (where the meanings are associated with different genders), Guaraní, Jarawara, Macaguán (by a semianalyzable term containing *-buk* ‘heel’), Rama (by a semianalyzable term containing an element meaning ‘tree’ and colexifying ‘grater’), Tehuelche, Yanomámi, Mandarin, White Hmong, Samoan (by the analyzable term *‘au-vae* ‘CLASS.LONG.OR.NARROW.THINGS-foot/leg,’ colexifying also ‘foot of hill’), Takia (by the analyzable term *awa-n to-n* ‘mouth-3SG arm-3SG’), Tetun, and Yay. Moreover, Katcha has *ku’ba mo buruŋe* ‘bone GEN chin,’ Waris *keu-mul* ‘chin-edge,’ Carib *-ekusali yepo* ‘-chin POSS,’ Cubeo *yedũ-cũ* ‘chin/jaw-CLASS.ROUND.ON.ONE.SIDE.AND.PLAIN.ON.THE.OTHER,’ and Hawaiian *papa ‘auwae* ‘flat.surface chin’ for ‘lower jaw’ specifically.

Alongside the complex or semianalyzable terms containing an element meaning ‘mouth,’ this association is also present in San Mateo del Mar Huave *mi-machat ombeayaran*, analyzable as ‘AL.POSS-machete mouth,’ and the situation in Baruya, with elements meaning ‘mouth’ and ‘bone,’ is precisely mirrored in San Lucas Quiaviní Zapotec and Bororo. Semianalyzable terms with an element meaning ‘mouth’ are alongside Kiliwa also found in Sora, Santiago Mexquititlan Otomí, and Xicotepec de Juárez Totonac. In Sora, moreover, *tʰo:dən ~ tʰo:dən ~ tʰu:dən* ‘jaw’ is colexified with ‘mouth of animal or vessel’ more specifically, and Muna *ghongki* can also refer to the “inner part of the mouth.” A lexical tie with

'bone' is, alongside the languages mentioned so far, also found in Tsafiki (*te'fún chide* 'tooth bone,' this term colexifies 'gums,' as is the case in Upper Chehalis) and Wichí, which has *toch'ay lhile*, containing *ch'a* 'cheek' and *lhile* 'bone.' A term with such structure is also found in Kiliwa, and there is a semianalyzable term in Carrier. As for 'tooth' as a constituent, Hawaiian has *papa niho* 'flat.surface tooth,' colexifying "row of teeth, set of teeth" as well as "bridge for false teeth," and as mentioned above, there are semianalyzable terms in Ket and Embera. Moreover, Ngambay and Wappo colexify 'jaw' with 'molar tooth' (and diachronic development from 'jaw' to 'back tooth' is attested in Romanian, Buck 1949: 221), and further analyzable terms of the lexical type for 'jaw' involving a constituent meaning 'tooth' are found in Hupda (*tag-cá?* 'tooth-box,' for which compare Spanish *quijada*, derived historically from Latin *capsa*, *capsus* 'box,' Buck 1949: 221), and Fijian (*kau-ni-bati* 'stick/wood/tree-poss-tooth'). Cubeo has a derived term.

The Kiliwa and Wichí association with 'cheek' is mirrored in Buli, Yoruba, Gurindji, Yir Yoront, Welsh, Lesser Antillean Creole French, Wintu, Abipón, Miskito, Kapingamarangi (by the analyzable term *gau wae* 'side foot/leg'), and Manange by colexification, and in Yir Yoront by the term *wal-kur* 'cheek/temple-slashing.sword.' Furthermore, Rendille *gi'dáamme* is grammatically the plural of *gi'dáam* 'cheek,' and in Basque, the meaning of *baraila* varies dialectally between 'jaw' and 'cheek' (it can also assume the meaning 'revelry, fray'), and for Hani *baqxoq* 'jaw' (perhaps related to *xoq* 'lock'), compare *baqba* 'cheek.'

Finally, Dongolese Nubian and Nez Perce colexify 'jaw' with 'palate' (the relevant Dongolese Nubian term is also the name of a village).

Other associations include: Efik *mbañ* colexifies 'jaw' with 'gill of fish' as well as, figuratively, "[l]oud, impertinent talking." Hausa *ma-k'aru* is analyzable as 'LOC-finish.' This term is primarily applied with reference to fish, and also colexifies "top, inside, of a cooking-pot" and other vessels. Noni *kediew* consists of *diew* 'word, language, noise, sound' and the noun class marker *ke-*. One *sesu* also means 'to squish.' Yir Yoront *thamanwalq* contains *man* 'throat, neck' and *walq* 'hollow place,' while Central Yup'ik *agluquq* is analyzable as /agluq-quq/ 'center.beam.of.a.structure-one.that.is.' Miskito colexifies 'jaw' with 'beard,' Toba with 'operculum,' and Hawaiian with 'cheekbone' as well as "to talk a lot, jabber" and culture-specific instruments. Lenakel *kauga* also means 'corner,' Tetun *hasan* also means 'ambush,' as well as "notch in wooden columns of buildings for the positioning of beams."

### 119. *The Kidneys*

Representation: 84%

Motivated: 24.6%

Thereof Analyzable: 17.1%

Thereof Colexifying: 7.5%

Thereof by Contiguity: 2.3%

Thereof by Similarity: 13.5%

Recurrent associated meanings: back/behind, heart, seed, fruit, liver, bean, testicle, stone, fat, urine/urinate

The 'kidneys' are often named by terms making reference to their particular shape. Common are comparisons with fruits of similar shape, either by colexification or by morpho-

logical complexity. In Ngambay, *mùnjù* also means 'bean' (a pattern of colexification shared with Kiowa, where it can also refer to a 'pea,' and mirrored by the complex Chaya-huita term *maquira-pi* 'bean-CLASS.FRUIT.BODY.OR.BACK'), in Mbum, *ánjòkò háì* consists of *án-jòkò* 'peanut' and *háì* 'flesh,' and in One, *wasou tala* contains *wasou* 'betel nut' and *tala* which either means 'seed' or 'round thing' more generally. In Kapingamarangi, the word for 'kidney,' *ibu mimi*, contains *ibu* 'coconut shell' (which also means 'cup,' presumably by functional contiguity) and *mimi* 'urine.' In Yuki, *pohot* ~ *pōhōt* 'kidney' may be identical with *pohut* ~ *po'hot* 'oak galls' and in Wappo, spoken in the vicinity and perhaps genealogically related, *ʔáwe* also may refer to potatoes or "any variety of food bulbs and corms." Fijian *ivi* is also the name of the native chestnut tree, and Nunggubuyu *wudu* is also the name of a tree with kidney-shaped nuts.

Frequent are also associations with 'fruit.' Kwoma has *magiir wuwu siik* 'back swell.up fruit/seed,' Toaripi *kōu uti fare* 'back bone fruit,' Hawaiian *kōnāhua*, perhaps containing *konā* 'hard' and *hua*, meaning 'fruit,' 'seed,' and 'egg' inter alia. (this term colexifies 'grease' and 'fat,' for this, compare Samoan *fatuga'o* 'heart/seed-fat/lard'), Tetun *fuan-sorin* 'heart/fruit-side,' and semianalyzable terms are found in Piro (where the relevant term also means 'seed' and acts as a diminutive marker; in Koyraboro Senni, there is a dialectal variant which is identical structurally with a verb meaning 'to sow, plant, raise'), Bwe Karen (where 'fruit' and 'round object' are colexified) and Lenakel (where this term colexifies 'side of ankle' and 'earlobe'); in Yay, the word for kidney, *maak<sup>2</sup> yaaw<sup>1</sup>*, includes the classifier for fruits, *maak<sup>2</sup>* and in Hani, *hhoqsiq* contains the classifier for round things *siq* (*hhoq* means 'to wear clothes, be covered' inter alia). As for 'seed,' Kwoma has *magiir wuwu siik* 'back swell.up fruit/seed,' One *wasou tala* 'betel.nut seed/round.thing,' Jarawara *siro noki/noko* ~ *kasiro noki/noko* 'frog/tree.species eye/face/seed/color' (note that there is a semianalyzable term with the identifiable constituent meaning 'eye' in Kildin Saami), Hawaiian *kōnāhua*, perhaps containing *konā* 'hard' and *hua*, meaning 'fruit,' 'seed,' and 'egg,' and Samoan *fatuga'o* 'heart/seed-fat/lard' (the constituents glossed as 'seed' have still other meanings in some languages).

Takia colexifies 'kidney' with 'stone,' Wichí has a derived term, Rama has *ngalíng kálup*, containing *ngalíng* 'stone' and *up* 'eye,' and Kyaka has a term where the identifiable constituent means 'hard' and 'pebble' alongside other things.

As evidenced by the One and Bwe Karen cases, at times 'round object' in general is associated with 'kidney.' This is also the case in Lengua, where a semianalyzable term exists with the identifiable constituent meaning 'to be round.' A term which also appears to make reference to the shape of the kidney by way of metaphor is found in Arabela, where *cajiniajajau* appears to contain *cajinia*, the word for a sitting person or a baby that is old enough to sit and the classifier *-jajau* for round objects. On a more abstract level, terms for the kidneys making reference to their shape are found in Abzakh Adyghe, which has *č'ent'əɫ°* /č'(e)-n-t'(e)-ɫ°(e)/ 'end-EPEN-poke.out/bend-pointed' and Pawnee, where *spiruusu<sup>2</sup>* probably contains underlying *as-* 'foot,' *piruus* 'crooked' and the nominal suffix *-u<sup>2</sup>*.

There are also languages in which the kidneys are conceptualized via their position in the body. This is found in two languages of New Guinea, Kwoma and Toaripi, as

seen above. In four languages in the sample, Hausa, Koyraboro Senni, Muna, and Abzakh Adyghe, ‘back’ and ‘kidney’ are associated by colexification (‘lower part of the back’ in particular in Muna). In Tetun, which has several words for the kidney, *fuansorin*, consisting of *fuan* ‘heart, fruit’ and *sorin* ‘side,’ is encountered. The Tetun example provides the transition of the discussion to connections between the ‘kidneys’ and other internal organs of the body. Most frequent are indeed those with the ‘heart,’ occurring in six sample languages, namely Ngambay, Guaraní, Yanomámi (where the relevant term may also be used to refer to other spherical objects) by colexification, and in Tetun, Samoan, and Imbabura Quechua (*yana shungu* ‘black heart’) by analyzable terms. Three languages, Badaga, Kiowa, and Ancash Quechua, colexify ‘kidney’ and ‘liver’ (Badaga also ‘lungs’ and ‘larynx,’ and Kiowa has several complex terms for both meanings available for disambiguation), and in Yanomámi, *amoyōri* also denotes the ‘gall bladder.’ A body-part metaphor is found in Khalkha and Rotuman, where ‘kidney’ and ‘testicle’ are colexified, while in Ancash Quechua, the ‘kidney’ is also called *ruru-n* ‘egg/testicle-3sg’ (see Jóhannesson 1949: 107 for Indo-European parallels). In Khoekhoe, *lnāb* is a semantically very general term that can be used to refer to ‘offals’ generally, including virtually all internal organs of the trunk.

Less frequent are terms for the kidney in which their function plays a role. Already mentioned was Kapingamarangi *ibu mimi* containing *mimi* ‘urine,’ in Sedang, *bong núm* appears to be analyzable as /bông núm/ ‘white urinate,’ and there is a semianalyzable term featuring an element meaning ‘urine’ in Toba.

Other associations include: Buli *yiini* also means ‘to live on others, be greedy,’ the Efik term *ek’put* is also used inter alia to refer to “the tassels on the neck of a goat,” and Hausa *k’oda* is also a verb meaning “[r]epair and sharpen the edge of a tool by beating.” The Burarra term *-gurday* is lexically connected to *gu-day* ‘friendship,’ and Dadibi *ene nawe* appears to contain *ene* ‘arrow, rattan.’ Meyah *ofómáf* appears to contain *ofóm*, which can mean ‘ripe’ or ‘root.’ Sahu *gogolatila* might contain *gogo* ‘body hair, fur, feather,’ and Sko *háng* also means ‘end of intestine.’ Yir Yoront *man-nerp* is analyzable as ‘neck-spirit.child,’ and *mortworrqworrr* contains *mort* ‘pile’ and *worrqo* ‘big.’ Kolyma Yukaghir *mumul* also means ‘fist,’ while the Cahuilla term *pípiviskun* is etymologizable as /pípivis-kunil/ ‘vomit-sack.’ Kiliwa *snpap* contains *pap* ‘bundle.’ The Yaqui term *sikupuriam* contains *siiku* ‘navel,’ while San Lucas Quiavini Zapotec *x:quèe’ts* also means ‘gizzard.’ Cavineña *epecaca* appears to be analyzable as /e-pere-caca/ ‘INAL.POSS-rib-small,’ and Cubeo *joedu* shares the root *joe* with a term denoting a tree species and an ‘axe,’ with different classifiers differentiating between the meanings. Hawaiian *pu’u-pa’a* is analyzable as ‘mound-firm.’ The term also means ‘virgin, virginity’ and is used figuratively to refer to emotions. Mandarin *yao1* also means ‘hip,’ Rotuman *ififi* also “in bunches or clusters,” and Yay *yaaw*<sup>1</sup> also ‘monster, ogre.’



120. *The Lip*

Representation: 94%

Motivated: 47.4%

Thereof Analyzable: 23.5%      Thereof Colexifying: 24.6%

Thereof by Contiguity: 16.1%      Thereof by Similarity: 8.9%

Recurrent associated meanings: mouth, edge, skin, beak, peel/rind/shell, bark, snout, female genitalia, entrance/door, language, cover, leather/hide, leaf

When words for the ‘lips’ are analyzable, the most frequent pattern is that the respective words consist of those for ‘mouth’ and ‘skin’ (with ‘skin’ at times having related meanings), as in San Lucas Quiaviní Zapotec *gui'dy ru'uh* /*gui'ihdy ru'uh*/ ‘skin mouth.’ This is found in nineteen of the sampled languages, next to San Lucas Quiaviní Zapotec in Buli (where an additional element meaning ‘leaf’ is present; the whole term also can refer to the “notch of a flute”), Efik, Mbum, Angkor, Kwoma, One, Toaripi, Yei, Sora, Highland Chontal, Kashaya, Lakhota, Xicotepec de Juárez Totonac, Cavineña, Cayapa, Kaingang, Maxakalí (where ‘skin’ colexifies ‘cover,’ compare also Yana *sob?liyawal(la)* /*sab?li-yauwal(la)*/ ‘cover-NMLZ-mouth’), and Yay, and thus occurs everywhere except Eurasia (the pattern is also etymologically detectable for Nunggubuyu); there is also an optional compound in Bislama of this type to disambiguate a term meaning both ‘lip’ and ‘mouth’ and semianalyzable terms are found in Sko, Haida, and Xicotepec de Juárez Totonac. Rama has *sík úk* ‘tooth skin’ (the language features also another semianalyzable term where the identifiable constituent is ‘skin’). Since ‘skin’ is lexically associated with several related meanings (see discussion in section 135), there are also secondary associations in many of the abovementioned terms. This is the case with ‘bark’ in Efik, Kwoma, Toaripi, Xicotepec de Juárez Totonac, Kaingang, and Maxakalí, with ‘rind,’ ‘peel,’ or ‘shell’ in Efik, Lakhota, Xicotepec de Juárez Totonac, Cayapa, Kaingang, Maxakalí, and Rama, with ‘leather’ in Sora and Cavineña, and with ‘cover’ in Lakhota. There is a semianalyzable term with the identifiable constituent colexifying ‘skin’ with ‘bark’ and ‘covering’ in Sko.

There is also a wealth of complex terms in other languages where an element meaning ‘mouth’ is present, but the other constituent does not mean ‘skin’: one relatively frequent variant of the complex terms of the ‘mouth-skin’ type is to have ‘mouth-edge,’ which is for instance found in Kiliwa (*ha?hiy* ‘lips,’ *ha?* ‘mouth’ *hiy* ‘edge, border;’ this term is itself glossed as ‘border;’ the language also has the term *ha?=nat-u?=kw+waa* ‘mouth=top/atop-OBL-WH+sit’ for the ‘upper lip’ specifically) as well as in Toaripi, Sora, and Fijian. When an association with ‘edge’ occurs, however, this is more frequent by colexification, which is found in sixteen languages, namely Ngambay, Rendille, Muna (also colexifying ‘side’), Sahu, Basque (also colexifying ‘corner’), Welsh, Itzaj, Bora, Bororo, Carib (where ‘upper lip’ is the colexified meaning more specifically), Cashinahua, Chayahuita, Embera, Guaraní, Miskito, Piro, Yanomámi, and Malagasy; furthermore, Yanomámi has *kasiki*, consisting of *kasi* ‘edge’ and the quantal classifier *ki* (see § 4.4.1.1.). In two languages, Buli and Samoan, terms for ‘lip’ are found which consist of the respective words for ‘mouth’ and ‘leaf;’ Hausa also has an association by colexification with ‘leaf bud.’ Further complex terms involving a constituent meaning ‘mouth’ are Efik *mfut inua* ‘shade mouth,’

Kanuri *kâ-cî-bè* ‘stick-mouth-of,’ Khoekhoe *nino a’dəgə* ‘mouth above’ for the ‘upper lip’ specifically, San Mateo del Mar Huave *apal ombeayaran* ‘lid mouth,’ Pawnee *haaka-huukita-haahiri*?, analyzable as ‘mouth-on.top-LOC,’ also for the ‘upper lip’ specifically, Yaqui *tenberia*, analyzable as /teeni-be’eri-a/ ‘mouth-exceed-NMLZ,’ Copainalá Zoque *ajsis /arɲaca-sis/* ‘mouth-flesh,’ Arabela *rupaa-que* ‘mouth-CLASS.CLOTH,’ Cubeo *jije-tarabw* ‘lip/mouth-CLASS.CIRCULAR,’ Kapingamarangi *malau ngudu* ‘two.things.joined.together mouth,’ and Lenakel has *nivhivhi-nhul* ‘end/tip-mouth’ for the ‘upper lip’ and *netpi-nhul* ‘belly-mouth-’ for the ‘lower lip.’ In Ngambay, Swahili, Burarra (where there is a very general term referring to the entire mouth area, including ‘cheek’ and ‘chin’), Kyaka, Acoma, Cheyenne, Comanche, Itzaj, Nez Perce, Nuuchahnulth, Oneida, Wintu, Bororo, Huambisa, Imbabura Quechua, Bislama, and Tetun, ‘lip’ and ‘mouth’ are directly colexified (Sahu colexifies ‘lip’ with “the outside of the mouth,” and Nez Perce also with ‘mouth of river’ and ‘cave;’ this association is frequent in Indo-European, Buck 1949: 229-230). Semianalyzable terms where one constituent is ‘mouth’ are found in Katcha, Khoekhoe, Abzakh Adyghe, Japanese, Kolyma Yukaghir, Biloxi, Chickasaw, Wappo, Aguaruna, Hupda, and Rotuman. Mostly due to colexification with ‘mouth,’ the ‘lips’ are also associated with ‘language,’ which is the case in Wintu, Imbabura Quechua, and Hawaiian; the only language including a term which colexifies ‘lip’ and ‘language,’ but not ‘mouth’ is Hawaiian.

In Swahili, Kyaka, Carrier, Central Yup’ik, Jarawara, Yanomámi, and Takia, ‘lip’ is colexified with ‘beak’ (see also section 5; in Kyaka with “wide beak,” in Central Yup’ik with “upper or lower part of a snout or beak” and in Takia with ‘pounting lips’ more specifically). The Pipil and Jarawara terms colexify ‘lip’ with ‘snout’ (in Jarawara alongside ‘fruit’ and ‘mass, hunk, lump, whole’), while in Acoma, ‘upper lip’ more specifically is colexified with ‘snout’ and conversely, in Central Yup’ik ‘upper part of snout’ is colexified with ‘beak.’ In four languages, Ngambay, Burarra, Kyaka, and Macaguán, the word for ‘lip’ is extended to also mean ‘entrance’ or ‘door;’ however only in Macaguán does the respective term not also denote the ‘mouth’ at the same time, a meaning for which this metaphorical extension is more common (see section 124 and § 6.2.3.2.). In four of the sampled languages, Ngaanyatjarra, Lake Miwok, Bislama, and Hawaiian, there are extensions to the female genitalia (as in Latin; in Hawaiian among other meanings colexified). Rarer metaphorical extensions are found in Nivkh, where *ɲapu* also means ‘bay,’ in Central Yup’ik, where *qerrluq* may refer to ‘stones around a firepit,’ and in Cashinahua, where *kebichi* (given as *kebixi* in the Spanish-Cashinahua section of the consulted source) may also be used with the meaning ‘eyelid.’ Finally, in five of the sampled languages, ‘upper lip’ and ‘lower lip’ are expressed by different lexical items. These languages are Toba, Lenakel, Dadibi, Lavukaleve, and Katcha.

Still other associations are: Efik *dum’baru* colexifies “[f]irst young leaf buds of any tree,” “[g]erminating of seeds,” and “[f]irst coming through of a tooth.” Kyaka *kambu* also means ‘vote, voice’ (likely due to Kyaka colexifying ‘lip’ with ‘mouth’), Meyah *ofúj ofóu* contains *ofóu* which can mean either ‘egg,’ ‘meaning,’ or ‘dust,’ and Waris *muemb-ta* appears to be analyzable as ‘saliva-small.object.’ Itzaj *chi* is also the name of a tree, while Nez Perce *tú’skiriñikw sipé’wn* ‘upper lip’ is analyzable as ‘uppermost twist-ACT.PTCPL.’ Central Yup’ik *qerrluq* also means ‘stones around a firepit,’ and Wayampi *εμε* also ‘kerf in flute,’

‘fabric edge of hammock rope,’ and ‘shore.’ Hani *meiqjil* contains *jil*, which inter alia means ‘to finish something’ (and for *meiq*, compare *meiqdaoq* ‘mouth, snout?’). Hawaiian *lehe* can also mean ‘stretched’ or ‘loose’ and also denotes a deep-sea fish.

#### 121. The Liver

Representation: 93%

Motivated: 13.4%

Thereof Analyzable: 1.5%

Thereof Colexifying: 11.9%

Thereof by Contiguity: 0.0%

Thereof by Similarity: 4.7%

Recurrent associated meanings: soul, heart, lungs, core/center/middle, belly/stomach, kidney, chest

There are relatively few lexico-semantic associations concerning the ‘liver;’ the ones that are attested link it mostly to other internal organs of the body. In five sampled languages, Ngambay, Kwoma, Sahu, San Lucas Quiaviní Zapotec, and Yay, colexification of ‘liver’ and ‘heart’ is found, in Imbabura Quechua, the ‘liver’ and the ‘kidneys’ are called *yana shunga* ‘black heart’ (in Irish, ‘liver’ is a compound of terms meaning ‘heavy’ and ‘heart,’ Buck 1949: 252). In Sahu, the simplex *katere* has rather broad reference, including, alongside ‘liver,’ ‘chest,’ ‘heart,’ and ‘soul’ (‘chest’ and ‘liver’ are also colexified in Efik). To single out the meaning ‘liver,’ *katere ma bibiwisi* may be used; *bibiwisi* is a reduplicated version of *biwisi* ‘banana flower,’ which has a reddish color similar to that of the liver. Other languages (Badaga, Kiowa, and Imbabura Quechua) colexify ‘liver’ and ‘kidney,’ and in still others, (Kwoma, Badaga, Laz, Cahuilla, and Yay), ‘liver’ and ‘lungs’ may be referred to by the same lexical item (this association may also be present in Angkor, though marked as dubious in the source). In the case of three languages, Ngambay, Khalkha, and Yanomámi, terms for ‘liver’ also include meanings such as ‘belly,’ ‘tummy,’ or ‘stomach.’ In Khoekhoe, there is a term covering internal organs (‘lungs,’ ‘heart,’ ‘liver,’ and ‘kidneys’) which is also capable of referring to the ‘stomach.’

Relatedly, in Central Yup’ik, *tenguk* also means ‘solar plexus,’ and in Badaga, *iraḷu* also denotes the ‘entrails’ and ‘intestines.’ In four languages of the Americas, Highland Chontal, Wappo, San Lucas Quiaviní Zapotec, and Yanomámi, the words for ‘liver’ also have a more abstract reading of ‘core,’ ‘center,’ or ‘middle,’ and Efik *ě’sět* inter alia also denotes the ‘inside, interior’ of something (Bowden 1992: 36, as cited by Heine and Kuteva 2002: 199, reports a grammaticalization path ‘liver’ > ‘in’ in a number of Oceanic languages).

The liver is also the seat of emotions or the soul in a number of languages (or more appropriately put, conceived of this way in the associated cultures). This is noted for the African languages Efik and Noni, and is also found in New Guinea in Kwoma, Meyah (see examples in § 4.5.1.4.1.), Sahu, Sko, and Toaripi, as well as in White Hmong in South-east Asia. In addition, in Fijian, ‘courage’ and ‘cowardice’ are associated with the ‘liver,’ Hawaiian *ake* is ambiguous between the nominal reading ‘liver’ and a verbal reading that has to do with desiring or yearning for something; optionally, *ake-pa’a* ‘liver-firm’ may be used to single out the nominal reading specifically, while in Embera, the sequence *dadyí* in

*dadyíthārí* ‘liver’ occurs in many expressions having to do with human beings, their souls and their spirit.

Other associations include: Hausa *hanta* is also used figuratively for “one’s dearest possession,” Koyraboro Senni *tasa* also denotes a ‘metal bowl’ and means ‘to push’ as a verb, while Ngambay *wùr* also means ‘patience.’ Baruya colexifies ‘sorrow,’ and Kyaka *pungi* also means ‘rope.’ One *wala* is also used with the meaning ‘side,’ Meyah *odóu* also means ‘front,’ and Sahu *katere* also ‘chest.’ Sko *pung* is also the name of a large bamboo species and also means ‘to butcher, cut up meat.’ Kosarek Yale *bubu* also means “carrying or bearing part, point of support.” Abzakh Adyghe *s<sup>o</sup>’a* also means ‘good, useful’ (though it occurs with the meaning ‘liver’ only in a redundant complex term). Badaga *cuttage ~suttagé* also denotes the ‘larynx,’ and Basque *gibel* also means ‘listlessness, lethargy.’ Khalkha *elige* figuratively also denotes the ‘breast’ and ‘blood relatives.’ The Sora term *ə’gare:n* is also used to denote “pulp of fruits, pith, kernels etc.,” and Welsh *iau* also means ‘yoke’ and ‘Thursday.’ Highland Chontal *łada* is also glossed as ‘center of being.’ Pipil (Cuisnahuat dialect) *-el-tapach* is analyzable as ‘-inside-shell,’ and Maxakalí *xupkūnāy* might contain *xup* ‘to suck, sip, hang down.’ Hani *caoq* also inter alia means ‘to contain, to be present inside,’ ‘a frightening place where spirits live,’ and acts as a classifier for clumps or clusters of things. Mandarin *gan1* also means ‘dry’ (in both readings reflecting Early Middle Chinese *kan*, Pulleyblank 1991: 102), and Rotuman *āfe* also ‘thousand’ inter alia. Takia *ate-* is also used with the meanings ‘palm of hand’ and ‘plain among hills,’ and White Hmong *siab* also means ‘high, tall.’ Lesser Antillean Creole French *fwa* also means ‘time, turn, occasion’ and ‘faith.’

## 122. The Lungs

Representation: 93%

Motivated: 25.4%

Thereof Analyzable: 16.5%

Thereof Colexifying: 8.9%

Thereof by Contiguity: 7.9%

Thereof by Similarity: 5.8%

Recurrent associated meanings: liver, light, heart, foam, chest, float, air

Words for the ‘lungs’ are often associated with other internal organs of the body lexically, most often with the ‘liver.’ Kwoma, Badaga, Laz, Cahuilla, and Yay colexify ‘lungs’ and ‘liver’ (Badaga also ‘kidney’ and ‘larynx;’ perhaps Anggor does, too, but this is marked with a question mark in the source), and there are complex terms in Ket (*mensij* /ben-sēŋ/ ‘apart-liver/innards’), Hupda (*hə wowōw* ‘liver wring.out’), Ancash Quechua (*yuraq ñatin* ‘white liver’), Bislama (*waet-leva* ‘white/bright-liver’), Fijian (*yate-vuso* ‘liver-foam’), Hawaiian (*ake-māmā* ‘liver-light,’ *ake-makani* ‘liver-wind,’ and *ake-pāhola* ‘liver spread’), Kapingamarangi (*ade di baahi* ‘liver ART side’), and Tetun (*aten-book* ‘liver-move’). Note also Yoruba *èdò-fóro* ‘liver-lung.’ There is a semianalyzable term in San Lucas Quiaviní Zapotec, where the identifiable constituent colexifies ‘liver’ and ‘heart,’ and it is with the latter organ that ‘lungs’ is also often associated in the sampled languages. Buin, Burarra (by the analyzable term *mun-molma* ‘CLASS.DOMESTIC-warmth’), Kwoma, and Yay have terms capable of referring to both organs (and they are considered the seat of emotions in Buin and Bu-

rarra). Guaraní has *ñe'ã vevúi* 'heart light' (though *vevúi* alone can refer to the 'lungs,' too), and Ancash and Imbabura Quechua a term literally translatable as 'white heart' (*yuraq shunqu* and *yurak shungu* respectively). Furthermore, there is a semianalyzable term with the identifiable constituent meaning 'heart, breast' in Kaingang. Moreover, in Khoekhoe, there is a general term for 'offals' which can refer to virtually all internal organs of the trunk, while there are other lexical items for each one specifically.

Almost all other recurrent lexico-semantic associations already occur in the terms mentioned so far, making reference both to the light weight and the spongy structure of the lungs that sets them apart from other internal organs of the body. Alongside Guaraní and Hawaiian, terms where their light weight is used for denomination are Yir Yoront *ngerr-lolt* 'belly-light.in.weight/hollow' (the source notes that this term is "[g]iven fairly consistently also for 'heart'"), Welsh *ysgyfaint* /*ysgafn-aint*/ 'light-ABSTR,' Chickasaw *ishoppaya*, analyzable as /im-shoppaya-/ 'DAT-be.light-NMLZ,' Arabela *pa namaca* 'our light,' and 'lungs' and '(be) light' or 'something light' are colexified also in Nuuchahnulth, Bora, and Samoan (note in addition the similarity between Tsafiki *que'fó* 'lungs' and *que'fún* 'light'). This association is also recoverable etymologically in German (and Germanic more broadly, Kluge 2002). Kluge has it that "die Lungen werden als 'die Leichten' bezeichnet, weil Lungen von Schlachttieren als einzige Innereien auf dem Wasser schwimmen" / "the lungs are called 'the light ones' because lungs of animals for slaughter as the only innards float on water," see also Jóhannesson (1949: 105) for Indo-European more broadly. The 'lungs' are colexified with 'foam' in Buin, Kaluli, and Toaripi, and a complex term is, alongside Fijian, found in Chayahuita *sa'poro* (/sa'po'-ro') 'foam-CLASS.PILE.' Similarly, Guaraní colexifies 'lungs' with 'be foamy, sudsy,' Jarawara has *hasa-bori* 'make.bubbles-container,' and Wintu *łoso* denotes both 'lungs' and 'foamy crest on waves.'

Consistent with Kluge's explanation, Nuuchahnulth also colexifies 'to float' and Central Yup'ik has *pugtaun* /*pugta-(u)n*/ 'float-device.for;' this term colexifies 'lungs' with 'float' (in the sense of 'raft') as well as 'buoy' and 'life-vest.' Similar to the association with 'wind' in Hawaiian mentioned earlier, the fact that the lungs are used for breathing is mirrored by associations with 'air' in Oneida (*yewelálakhwa?*, analyzable as /ye-wel-l-hkw-a?/ 'FEM.INDEF.SG.AGENT-air/wind-be.in-INSTR-HAB') and Wichí, which has a very similar term; likewise, the meanings 'lung' and 'breath' are connected diachronically in Greek. Tasmanian (Middle-Eastern, Southeastern, and Western) and Lesser Antillean Creole French colexify 'lungs' with 'chest.' There are also associations with 'meat' and/or 'flesh,' however, only by semianalyzable terms, namely in Mbum, Abzakh Adyghe, and the Norton Sound - Upaliq dialect of Central Yup'ik.

Other associations include: Hausa *huhu* is also the name of a lung-related disease of horses (and has still further meanings). Koyraboro Senni *kumbu* also dialectally denotes a 'type of hoe,' and Baruya *pawajikawaai* ~ *pawajikawaalo* contains *pawajika* 'spleen.' Buin *turupa* also means 'saliva' (compare the association between 'foam' and 'saliva' described in section 24; Buin is among the languages colexifying 'lungs' with 'foam'). Dadibi *ogwa bai*, also meaning 'spirit,' is analyzable as 'son baby,' Muna *ghau* also denotes "overgrown tubers/maize." Sahu *aba'abala* appears to be reduplicated from *abala*, meaning "(wall made of) midribs of sago palm fonds." Badaga *cuttāge* ~ *suttāge* also is used for 'kidney,' and Haida

*hlgu* also denotes a 'fish's swim bladder.' Itzaj *b'oj* also means 'to knock' inter alia. The Kiliwa term *x-?+kwiiy-waa-u?* is semianalyzable as 'CAUS-DN+??-sit-OBL' (compare *-kwiiy* 'cloud?'), and Kiowa *k'ahyoudl* is related to *k'æ* 'skin' and *-houdl* 'intensive.' Wintu *χOs* 'fog, steam, gas, lungs' yields *χoso* 'lungs.' Copainalá Zoque *se'u* also means 'to have a sore throat.' Cayapa *jenana* might contain *nana* 'balsa' (notably, a particularly light kind of wood), and Cubeo *vũibo* consists of *vũi* 'tuber' and the classifier for round or hard objects *-bo*. Guaraní *tajygue* also means 'tendon, vein, nerve, muscle' and 'strength,' Kaingang *fe kãnhvy* also means 'joy,' while the Rama term *ikúngkungma* appears to be based on *kung* 'louse, whit, air root.' Toba colexifies 'lung' with 'tears,' Wayampi *tuluwiy* contains *tulu* 'big,' and Yanomámi *hereki ~ hērēki ~ heremopi* contains *here* 'wet' and a quantal classifier (see § 4.4.1.1.). Finally, Hani *povq* also means 'to get soft' and 'to be empty, used up,' Bwe Karen *θá'ó* also means 'to weed,' and Rotuman *ma'ma'a* is reduplicated from *ma'a*, which means, inter alia, 'light in color.'

### 123. The Milk

Representation: 91%

Motivated: 54.7%

Thereof Analyzable: 23.9%

Thereof Colexifying: 30.8%

Thereof by Contiguity: 48.9%

Thereof by Similarity: 2.4%

Recurrent associated meanings: breast, water/liquid/juice, udder/teat, nipple/teat, suck, resin, semen, pap, Burton's legless lizard, cow

There is one cross-linguistically dominant association in terms for 'milk' that is unsurprising, namely that with 'breast.' This is frequently realized by analyzable terms (40 languages), in which case 'milk' is canonically literally 'breast water' (e.g. Sentani *nimə bu*) or, more generally 'breast liquid' (e.g. Bororo *mokuro kuru*) or 'breast juice' (e.g. Kaluli *bo ib /bó ib/*). Such terms are also found in Efik, Muna, Waris, Ket, Kiliwa, Kiowa, Pawnee, Cavineña, Embera, Guaraní, Hupda, Jarawara, Lengua, Miskito, Piro, Tsafiki, Wayampi, Yanomámi, Great Andamanese, Hani (where *aqqul* colexifies 'sweet' and indeed likely also contains *qul* 'sweet'), Bwe Karen, and Malagasy; Buli and Kapingamarangi have semianalyzable terms where the identifiable constituent means 'breast.' In languages with systems of nominal classification, as already seen in § 4.4.1, the respective terms typically involve a classifier affix, as is the case in Arabela (where 'breast' and 'teat' are colexified), Bora, and Chayahuita; similarly, in Swahili, the same root yields both meanings depending on noun class assignment. Variants of the morphologically complex terms mentioned above are Meyah *méngk ofód* 'breast flood' and San Mateo del Mar Huave *aonts mijiwaran* 'excrete breast,' and there are semianalyzable terms with 'water' or 'juice' in Cayapa, Rama, and Great Andamanese.

A further variant are terms in which the constituent meaning 'breast' is replaced by a verbal element meaning 'to suck,' such as in Chickasaw *pishokchi*, which is analyzable as */pishi okchi/* 'suck juice/liquid.' The same pattern is also found in Rotokas (*roroo ovi* 'develop.breasts/breastfeed/suckle CLASS.LIQUID') and Tetun (*susu-been* 'suck-liquid'); however, Tetun also has a term of the derived type, namely *susu-n* 'suck-SINGULATIVE,' and in-

deed, derived terms with a similar structure are found in more sampled languages, namely in Khoekhoe, Kolyma Yukaghir, and Haida, and in Bislama and Samoan, *susu* is ambiguous as to part of speech and means 'milk' in nominal and 'to suck' in verbal usage. Similarly, Pawnee colexifies 'breast' with 'to suck,' and hence the relevant complex term for 'milk' also betrays an association with the latter meaning. Fijian *sucu* colexifies 'to be born, to suck the breast' with 'birth, milk' (there is an optional complex term with *wai* 'liquid' on the basis of *sucu*); the association with sucking is also attested weakly in Indo-European, the evidence being an Albanian word for 'cheese' containing a root attested in Sanskrit with the meaning 'suck' (Buck 1949: 385).

Colexification of 'milk' and 'breast' is even more frequent, meaning effectively that forty-nine languages have a single term to cover both 'milk' and '(female) breast': these are Ngambay, Noni, Yoruba, Anggor, Baruya, Buin, Burarra, Gurindji, Kyaka, Lavukaleve, Ngaanyatjarra, Nunggubuyu, Toaripi (colexifying also 'scrotum' and denoting a particular breast-shaped shellfish), Sahu, Kosarek Yale, Basque, Kolyma Yukaghir, Carrier, Upper Chehalis, Cheyenne, Haida, Kiowa (colexifying also 'to flow, melt'), Lake Miwok, Nez Perce, Nuuchahnulth, Oneida, Santiago Mexquititlan Otomí, Quileute, Xicotepec de Juárez Totonac, Wappo, Wintu, Yuki, Central Yup'ik (Yukon and Norton Sound dialects), San Lucas Quiavini Zapotec, Aguaruna, Aymara, Cashinahua, Huambisa, Kaingang, Macaguán, Bislama, Hani, Hawaiian (here there is the term *wai-ū* 'water-breast' which however can refer to 'breast' itself), Lenakel, White Hmong, Rotuman, Samoan, Takia, and Tetun. Biloxi may be another case of colexification. Here the source gives *wa'k tasi'* /*wak tasi'* / 'cow female.breasts' as the term for milk (for which compare Maxakalí *mūnūytut yōktat hep* contains *mūnūytut* 'cow' and *hep* 'liquid'), which suggests that this in fact refers to milk as a foodstuff and in fact *tasi'* alone can refer to both 'milk' and 'breast.' Hausa, by another term, colexifies 'breast' and 'sour milk' *inter alia*. For Kashaya, it is noted that *šĩ?do* 'breast' is also the old term for milk (for which there is now *molokko*, a loan from Russian). Interesting in this context is the case of Wintu, which features a single term for both referents but also has another term, *wuh*, which denotes both 'cattle' and 'milk,' and this kind of provenience contiguity may have come into being analogously.

Associations with 'breast,' either by analyzability or colexification, are common all over the world, with the notable exception of Eurasia, in which the association is only found in Basque and Ket; it is also not reported by Buck (1949: 385) for Indo-European. Redundant (or seemingly redundant) terms are also common, such as Yir Yoront *thayn-polqor* 'breast-milk,' although such formations may either be motivated by the introduction of cow's milk as a foodstuff and serve to distinguish mother's milk from it or to single out readings of the simplex for 'milk' if this is itself colexifying, for instance with other liquids. For instance, Kwoma *pi* ranges semantically over 'blood,' 'sap,' and 'milk,' and this may motivate the presence of *muku pi* (*muku* is 'breast').

Further patterns of colexification, such as that with 'blood' and 'resin' in Kwoma (the association with 'resin' is also found in Sora and Ineseño Chumash) already mentioned above include that with 'milky-looking pus' in Kyaka, that with 'soup' in Rotokas, that with 'milky juice' and 'latex' in Basque, that with 'juice' in Sora and Ineseño Chumash,

that with ‘semen’ in San Lucas Quiaviní Zapotec and Bislama, and that with ‘pap’ in Yoruba and Tuscarora.

A further pattern of colexification, encountered in Yoruba, Basque, Kolyma Yukaghir, Lavukaleve, Ineseño Chumash, Haida, Lake Miwok, Nuuchahnulth, Wintu, Aguaruna, Aymara, is that with ‘nipple, teat,’ and it is likely due to the colexification with ‘breast’ and ‘nipple’ by spatial contiguity rather than a genuine association, although Nez Perce and Ineseño Chumash colexify ‘milk’ with ‘nipple’ while having a separate term ‘for breast.’ In Cubeo, the ‘milk’-word *opeco* appears to be derived by unknown means from *ope* ‘nipple,’ though this remains somewhat unclear. The association with ‘nipple, teat’ is realized by analyzable terms in Arabela, Cavineña, Embera, Miskito, Wayampi, and Malagasy.

‘Milk’ is also colexified with ‘udder, teat’ in Hausa, Yoruba, Kyaka, Toaripi, Haida, Lake Miwok, Nez Perce, Nuuchahnulth, Tuscarora, Wintu, Aguaruna, Aymara, White Hmong, Rotuman, and Takia (which also colexifies ‘to squeeze’), and associated with these meanings due to ‘udder, teat’ being a secondary meaning of ‘breast’ by analyzable terms in Efik, Arabela, Cavineña, Embera, Miskito, Wayampi, Yanomámi, and Malagasy.

Both Burarra and Nunggubuyu extend their respective terms for ‘breast, milk’ to also denote “Burton’s legless lizard,” a species of lizard native to Australia and New Guinea.

Still other associations include: Hausa *madara* is also the name of a “kind of native-made cloth,” “English salt in loose form,” “[t]he unboiled juice of certain fruits or of sugar-cane,” “[u]nadulterated musc scent,” as well as “[p]ure silver.” Koyraboro Senni *waa* also means ‘to defecate, shit,’ and Yoruba *omú* is also the name of an “instrument used in weaving to divide the woof.” Muna *susu* ‘canned milk’ (borrowed in this sense from Bahasa Indonesia) is also the name of a large shellfish inter alia. Cheyenne *nénééhe* also means ‘bottle’ (this term is register-specific). Lesser Antillean Creole French *let* also means ‘letter, character,’ Nuuchahnulth *?inma* also denotes “Nob Point, where white powder seems to run out of a breast-shaped rock,” Wintu *?Em* also means “hold pectorally, carry something in the arms, embrace,” Bislama *titi* and *susu* also mean ‘unweaned,’ Hawaiian *waiū* also is the term for a ‘wet nurse’ and *kea* also means ‘white, clear’ inter alia. Rotuman *susu* also means ‘to sew.’

#### 124. *The Mouth*

Representation: 97%

Motivated: 48.8%

Thereof Analyzable: 5.9%

Thereof Colexifying: 42.9%

Thereof by Contiguity: 15.3%

Thereof by Similarity: 6.6%

Recurrent associated meanings: opening, word/language/speech, beak, lip, edge/tip, door/entrance, muzzle/snout, estuary, tooth, hole, blade, riverbank, barrel of gun/muzzle of gun

Ngambay, Swahili, Burarra, Kyaka, Acoma, Cheyenne, Comanche, Itzaj, Nez Perce, Nuuchahnulth, Oneida, Wintu, Bororo, Huambisa, Imbabura Quechua, Bislama, and Tetun colexify ‘mouth’ with ‘lip’ (a common association by semantic shift in Indo-European,



Buck 1949: 228); furthermore, Aguirre Licht (1999: 99) mentions *í-tǎe* ‘lip-intersection’ for Embera, but it is not entirely clear whether this term has lexical status, and Cubeo has a single root occurring in terms for both ‘lip’ and ‘mouth,’ with classifiers differentiating between the two.

Baruya, Gurindji, and Cahuilla colexify ‘mouth’ with ‘tooth’ or ‘teeth,’ and Jarawara has *inohoti/inohoti*, consisting of *ini/ino* ‘tooth’ and *hoti/hoto-ne* ‘hole’ (compare also Kosarek Yale *sikaan* ‘mouth’ and *si* ‘tooth, thorn, point, name’). In fact, similar terms are found in other languages of South America: Maxakalí has *yĩy-kox* ‘speak-hole’ and Tsafiki *fi’quí foró* ‘language opening/hole’ (though note also that Lesser Antillean Creole French colexifies ‘mouth’ with ‘hole’ among other meanings directly). Furthermore, an element meaning ‘tooth’ may be diachronically detectable in the word for ‘mouth’ in Nunggubuyu. Similarly, ‘mouth’ is colexified with ‘language,’ ‘word’ and/or ‘speech’ in Efik, Hausa, Ngambay, Rendille, Muna, Ngaanyatjarra, Khalkha, San Mateo del Mar Huave, Kiliwa, Wintu, Arabela, Bora, Chayahuita, Guaraní, Miskito, Ancash and Imbabura Quechua, Yanomámi, and Sedang, which also colexifies ‘knife’ (and Baruya colexifies ‘angry speech, anger’ more specifically), and in Tasmanian (Middle-Eastern, Southeastern, and Western) terms for ‘mouth’ are clearly related to those for ‘language, utterance, speak,’ though the precise relationship is not recoverable. Consistent with the South American pattern in complex terms just mentioned, the Cubeo term for ‘mouth’ is differentiated from ‘lip,’ with which it shares its root, by a classifier for entities having to do with language.

There are, however, also many metaphor-driven extensions of ‘mouth’ to other meanings. Buli, Khoekhoe, Laz, Sora, Lesser Antillean Creole French, Itzaj, Yanomámi, and Mandarin colexify ‘mouth’ with ‘muzzle’ or ‘snout’ (a common pattern in Indo-European, Buck 1949: 228), Buli, Efik, Khoekhoe, Ngambay, Swahili, Baruya, Kyaka (colexifying ‘wide beak’ specifically), Toaripi, Abzakh Adyghe, Nivkh, Kashaya (where *haʔbo* is analyzable as */ʔaha-ʔbo/* ‘mouth-enlarge/swell.up’ and means both ‘protrusion of the mouth’ and ‘external mouth’), Yaqui, Hupda, Fijian, Malagasy, Rotuman, Samoan, and Tetun with ‘beak’ (in Kyaka ‘wide beak’ specifically; similarly, Yir Yoront colexifies “bottom of bird’s beak” *inter alia*), and Buli, Hausa, Kanuri, Khoekhoe, Koyraboro Senni, Baruya, Basque, Upper Chehalis, Itzaj, Santiago Mexquititlan Otomí, Pipil, San Lucas Quiaviní Zapotec, Arabela, and Bororo colexify ‘mouth’ with ‘edge’ and/or ‘tip.’ Similarly, Welsh colexifies ‘mouth’ with ‘end’ (and also with ‘top’ and ‘head’).

Efik, Khoekhoe, Toaripi, Abzakh Adyghe, Upper Chehalis, Nez Perce, and Mandarin colexify ‘mouth’ with ‘estuary’ (compare the derivational relationship between Latin *ōs* ‘mouth’ and *ōstium* ‘door, entrance, river-mouth,’ as well as cognates of *ōs* in Old Prussian and Old Norse meaning ‘river-mouth’ according to Buck 1949: 228, and section 20 for evidence for this connection from the present sample). Nez Perce also colexifies ‘mouth of cave.’ Similarly, Efik *inua* also denotes an ‘inlet, gap,’ Lavukaleve *leu* also a “passage, channel in reef for a canoe to go,” and Buli and Tetun colexify ‘mouth’ with ‘riverbank.’

Buli, Efik, Hausa, Noni, Burarra, Kyaka, Ngaanyatjarra, Rotokas, Basque, Lengua, and Mandarin colexify ‘mouth’ with ‘door’ and/or ‘entrance’ (compare the Latin evidence quoted from Buck above). Dongolese Nubian, Rendille, and Basque colexify ‘mouth’ with

‘blade,’ and more specifically, Buin colexifies ‘cutting edge of blade,’ while Lake Miwok and Hawaiian colexify ‘barrel of gun’ or ‘muzzle of gun.’

Furthermore, ‘mouth’ is extended to ‘opening’ in general (but often of bottles or other vessels in particular) in Efik, Hausa, Khoekhoe, Ngambay, Dongolese Nubian, Yoruba, Kyaka, Muna, Ngaanyatjarra, Rotokas, Toaripi, Abzakh Adyghe, Basque, Khalkha, Sora, Welsh, Lesser Antillean Creole French, San Lucas Quiaviní Zapotec, Bororo (also with ‘curve’), Lengua, Miskito, Toba, Yanomámi, Bislama, Hawaiian, Lenakel, Mandarin, Rotuman, and Takia.

Other associations include: Buli *noai* has additional readings as ‘advice,’ ‘command,’ ‘oath,’ and others, while *takabi* means ‘sherd, piece of clay vessel’ and is a vulgar and insulting term for ‘mouth.’ Efik *inua*, similarly, colexifies ‘mouth’ with ‘report’ and ‘boasting’ inter alia. Hausa *baka* also means ‘bow’ and ‘catch of lock,’ again inter alia, and *baki* is glossed also as “conclusion, maturity.” Koyraboro Senni *mee* ~ *miñe* also means ‘loose’ inter alia and Ngambay colexifies ‘suck’ and ‘embrace,’ again among other meanings. Kyaka *kambu* also means ‘vote, voice.’ The Rotokas term *gisipo* also means ‘talk, words,’ while *akuta* also means ‘to open one’s mouth, to shout’ as a verb. Abzakh Adyghe *ʔə* also means ‘to speak,’ ‘beg’ as well as ‘to listen, understand’ and other things, Basque *aho* also means “articulation, diction” among other meanings, while Laz *p’ici* also means ‘face, front side’ and similarly, San Lucas Quiaviní Zapotec *ru’uh* also ‘in front of’ inter alia. Abzakh Adyghe *ze* also means ‘to grill, fry’ and ‘to wait’ and perhaps ‘knee’ and ‘physical deformation,’ Ket *qō* also ‘ice’ and ‘ten,’ while Khalkha *ama(n)* is also used as a unit for counting persons in census or as food-consumers. Sora *t’o:dan* ~ *t’o:dan* ~ *t’u:dan* may also refer to the ‘jaw’ (though from the source it is not clear whether the term assumes this meaning on its own or only in compounds). Haida colexifies ‘mouth’ with ‘to feed,’ and the relevant Itzaj term also denotes a tree species. Lakhota *í* also means ‘to arrive at a place away from here,’ and Santiago Mexquititlan Otomí also ‘to love.’ Abipón *n-aag-Rat* is analyzable as ‘POSS.INDEF/3SG-bite-CAUS,’ the association with ‘bite’ (as well as ‘lick’) is also recoverable for Nunggubuyu. Pawnee *haaka’u*’ also means ‘drum,’ while Yana *bal(la)* is also used dialectally with the meaning ‘cheek’ (an association evidenced by semantic shift in Indo-European, Buck 1949: 228). Kaingang *jēnky* appears to be analyzable as ‘eat-cut/smell.’ Miskito colexifies ‘mouth’ with ‘space, center,’ ‘road,’ ‘mandate,’ ‘message,’ ‘will,’ and ‘quarter,’ Wayampi *yulu* also denotes the mouthpiece of a flute, and Wichí colexifies ‘mouth’ with ‘hunger.’ Bislama *maot* also means ‘vagina’ in informal use and has a further meaning related to the preparation of Kava. Kapingamarangi *ngudu* also means ‘to melt’ as a verb. Hawaiian colexifies ‘mouth’ with ‘talky person,’ ‘neck of dress,’ ‘to carry on back,’ ‘load carried on back,’ among still other meanings. Rotuman *nuju* also can refer to a ‘spokesperson’ and “the most easily pierced in the three ‘eyes’ of a coconut-shell” inter alia.

125. *The Phlegm*

Representation: 49%

Motivated: 51.3%

Thereof Analyzable: 12.2%

Thereof Polysemous: 39.1%

Thereof by Contiguity: 6.6%

Thereof by Similarity: 15.0%

Recurrent associated meanings: saliva/spittle, cold/flu, snot, cough, pus, resin, semen, throat

Eight sampled languages, Hausa, Noni, Kyaka, Ngaanyatjarra, Jarawara, Miskito, Tehuelche, and Toba colexify 'phlegm' (or 'sputum') with 'flu' or 'cold,' while the same lexical item in Embera yields the two meanings depending on gender. Fifteen languages, Khoekhoe, Ngambay, Berik, Badaga, Khalkha, Haida, Quileute, Aguaruna, Arabela, Cashinahua, Cavineña, Cubeo, Guaraní, Huambisa, Ancash Quechua, Hani, and Hawaiian, with 'saliva, spittle' (Khalkha also with 'tears'), and in Embera, again gender is used as a device to differentiate the meanings. Hausa, Ngaanyatjarra, Badaga, Basque, Itzaj, Pawnee, Tuscarora, Embera, Tehuelche, and Fijian colexify 'snot.' Kaluli and Wintu colexify 'phlegm' with 'cough,' Muna with "cough up and spit out phlegm," Toba has a term for 'phlegm' derived from 'to cough,' Bislama has *doti blong kof* 'rubbish/pus poss cough/cold' (compare colexification of 'phlegm' and 'pus' in Badaga and Wintu), Samoan *fatu-tale* is analyzable as 'heart-cough,' and a semianalyzable term where the identifiable constituent is a verb meaning 'to cough up' is found in Kashaya. Chayahuita has *iro ipirin-so* 'cough/flu spit.out-3SG.SUB.' Furthermore, a semianalyzable term involving a constituent meaning 'nose' is found in Yir Yoront. Pawnee *kitutkuutu* is analyzable as /kitut-kuu?at-u?/ 'throat-rotten-NOM,' and Lesser Antillean Creole French has the (suspiciously long) term *sistans épé ki ka sòti an né ében gòj* 'substance thick REL PROG get.out in nose or throat;' there is a semianalyzable term containing an element meaning 'neck, throat' in Highland Chontal. San Lucas Quiaviní Zapotec colexifies 'phlegm' with 'semen,' and in Burarra, *gungulol* 'phlegm' is derived from *gulol* 'semen, rotten.' Hawaiian *hākelo* also denotes 'sticky sap from trees' (there is also another term, *male*, which has other seemingly unrelated meanings such as 'to marry' and also denotes a fish species), and in Toba, 'phlegm' and 'resin' generally are colexified.

Other associations include: Hausa *kaki* colexifies 'phlegm' with 'beeswax' among other meanings, and *majina* also means 'to blow the nose' inter alia when used verbally. Ngambay *wenren* also means 'vomit,' while Kwoma *ukwa siik* contains *siik* 'fruit.' Mali *genaing* also means 'cerebrum,' and Kosarek Yale *sikna* also 'dirt, bits of dirt.' Nivkh *čjevčjevu mif* is analyzable as 'be.wet soil.' Lake Miwok *lée?kaṭi* also has the verbal readings of "to cough up phlegm" and for phlegm "to come up by itself in the process of clearing one's throat," Carib *kòwe* also denotes the 'round-worm,' and Rama *síri* seems to consist of the elements *sí* and *ri*, which both make reference to 'water.' Muna *ngallangalla* also means 'phloem.' In Hani, *zaoqpeiv* is 'phlegm, spittle;' both *zaoq* and *peiv* have meanings on their own that do not seem to bear a semantic relationship to either 'phlegm' or 'spittle,' the closest among the meanings of *peiv* is 'to have a sickness or disease that continues, to have

a relapse of sickness' and its function as a classifier for dewdrops. Finally, Lenakel *noua-nelpwa* is analyzable as 'fruit-fat/grease.'

#### 126. *The Navel*

Representation: 91%

Motivated: 28.1%

Thereof Analyzable: 8.3%

Thereof Colexifying: 20.5%

Thereof by Contiguity: 17.2%

Thereof by Similarity: 5.5%

Recurrent associated meanings: umbilical cord, center, placenta/afterbirth, belly/stomach, family/relatives, cause/origin

The most common lexico-semantic association for this meaning is that with 'umbilical cord.' It is realized by colexification in Buli (where the term also colexifies 'stalk of calabash' inter alia), Kwoma, Lavukaleve (where the term also means 'moonshell'), Meyah, Rotokas, Toaripi, Sahu, Sko (the relevant term is *kóengri*, for which compare *kóeng* 'tooth'), Waris, Kosarek Yale, Yir Yoront, Khalkha, Cheyenne, Ineseño Chumash, Haida (where the term also colexifies "mouth of sea urchin"), Nuuchahnulth, Quileute, Yana, San Lucas Quiaviní Zapotec, Aymara, Carib, Bislama, Hawaiian, Lenakel, Manange, and Rotuman. In Cubeo, the same root suffixed with different classifiers yields the meanings 'navel,' 'umbilical cord,' and 'liana' respectively. Miskito has *tukta awa* 'child ribbon,' suggesting that 'navel' and 'umbilical cord' are also colexified here. Furthermore, there is an analyzable term in Kashaya (*?ohqomo* /*?ohqo-?imo*/ 'umbilical.cord-hole'), and a semianalyzable term in Bwe Karen; in San Lucas Quiaviní Zapotec, the relevant term also may refer to a "hole in a container of liquid or a small dam that can be closed with a plug or stopper." 'Navel' is colexified with 'stomach' in Basque, and dialectally, the meaning of Kyaka *kumu* varies between 'abdomen' and 'navel' (the term also denotes a kind of tree), and complex terms betraying this association are featured in Toaripi (*ére lalave* 'belly corner'), Abzakh Adyghe (*nəbepc'eəy*, containing *nəbe* 'belly, stomach' and *pc'e* 'false'), and Cayapa (*ajbundyu*, presumably analyzable as /*ajca-bundyu*/ 'abdomen-knot'). Kwoma, Waris, Cheyenne, and Wintu colexify 'navel' with 'placenta' or 'afterbirth,' and Kyaka has *reme lyolo* 'placenta end.'

Moreover, seven sampled languages, Hausa, Kyaka (by the term *mumbi renye*, which appears to contain *mumbi* 'kind of reed' and *renye*, meaning 'source, origin, reason' alongside other things, for which compare Efik *i'buüt nsia*, analyzable as /*i'buüt nsi'a*/ 'head/top/cause/origin bowels'), Khalkha, Welsh, Guaraní, Hawaiian, and Malagasy extend terms for 'navel' to the abstract, namely to 'center' or like meanings (see Buck 1949: 248 for Indo-European data; Welsh in addition also colexifies 'heart'). Hausa, furthermore, also colexifies "central supporting pillar of a ceiling" and 'birthplace' inter alia, and Sedang, similarly, denotes the "center of *chúa tô'u* weaving pattern" inter alia, and Yay *dud* also means 'center of flower.' Finally, in Khalkha, Wintu, and Hawaiian, relevant terms can also refer to the 'family' or 'relatives.'

Other associations include: Buli *siuk* also means 'street, road' inter alia. Khoekhoe *sunis* is presumably related to *sun* 'to joint, cut meat at joint,' (similarly, Wintu *naq* might

be related to *nEk* meaning ‘cut off’ inter alia; *naq* also means ‘to pity, care for, feel sentiment, empathy towards’), and Hausa *cibi* is also the name for a small metal spoon as well as for ‘umbilical hernia.’ Buin *ituge* also denotes the “anal vent in fish,” Burarra *jirrcha* also the tuber of the water lily *Nymphaea violacea*, and Kaluli *ho:nduf* contains *ho:n* ‘water.’ Kwoma *nawaba* also means ‘stem of plant’ and ‘trunk of tree,’ ‘carcass,’ as well as “neck of a slitdrum.” Ngaanyatjarra *nyuntjirn(pa)* also means “spinning through the air very fast (of spear),” while Nunggubuyu *mala* also means ‘honey’ and ‘(clear) sky.’ The Southeastern Tasmanian term *tūnā* is also the name of the ‘triton shell,’ while Badaga colexifies ‘navel’ with ‘umbilical rupture,’ Khalkha with ‘isthmus’ inter alia, and Welsh with ‘nave of wheel’ and ‘boss of shield.’ Chickasaw *ittialbish* ~ *ittihalbis* may contain *iti* ‘mouth,’ and Itzaj *tuch* also denotes the ‘gizzard of birds.’ Kiowa *p’eip’out* may contain *p’ei* ‘vulva,’ while Pipil *-xi:k* also means ‘bell clapper’ in the dialect of Santo Domingo de Guzmán, and Tuscarora *u?nihsèreh* also means ‘star, planet.’ Wintu *bō’s*, derived from *bō* ‘live, reside, remain, keep; stay, be in a sitting position, sit, dwell, stay,’ also means ‘home, camp,’ and Central Yup’ik *qallaciq* is perhaps analyzable as /qalla-(u)ciq/ ‘be.boiling-condition.of.possessor.with.respect.to,’ with the presence of the postbase (see § 4.4.2.) not being entirely clear. Chayahuita *imëra* seems to be analyzable as /imërin-ra/ ‘smell-CLASS.SMALL.THINGS,’ and Kaingang *nūg-nin* as ‘twist-tuber/potato.’ Rama has a semianalyzable term containing an element meaning ‘eye,’ and Tehuelche *wet’* is also one of the variants of *wet’* ~ *wet* ~ *wet’e* ‘to eat something.’ For Fijian *vicovico* compare *vico*, the name of a variety of wild sugar cane. Kapingamarangi *bida* is also used with the meaning ‘end’ and ‘piece,’ and Mandarin *qi2* also means ‘to ride’ (reflecting different Early Middle Chinese etyma, Pulleyblank 1991: 246). Finally, Rotuman *pufa* also denotes a “shallow fire-hole for open fire, fire-place.”

#### 127. *The Neck*

Representation: 96%

Motivated: 39.9%

Thereof Analyzable: 2.9%

Thereof Colexifying: 36.9%

Thereof by Contiguity: 30.3%

Thereof by Similarity: 4.7%

Recurrent associated meanings: throat, nape, voice, collar, neck of vessel, top, behind, Adam’s Apple, top, rain, swallow, head

Many sampled languages colexify ‘neck’ with ‘throat’ (see also Buck 1949: 231-232 for Indo-European data). This is the case in Efik, Berik, Burarra (also with ‘larynx’), Gurindji, Ngaanyatjarra (where the relevant term denotes the front of the neck specifically), Muna, Sentani, Toaripi, Yir Yoront (where the relevant term may also refer to the “upper chest area,” “mid-trunk area,” and has metaphorical uses), Greek, Kildin Saami, Welsh, Highland Chontal, San Mateo del Mar Huave, Itzaj, Kiliwa, Kiowa, Wintu, Yaqui, Yuki, San Lucas Quiaviní Zapotec, Copainalá Zoque, Embera, Lengua, Ancash Quechua, Tehuelche, Tsafiki, Hawaiian (by the formally redundant term *kani-‘āī* ‘hard-neck’), Bwe Karen, Mandarin, White Hmong, Rotuman, and Tetun (the relevant Rotuman term also means “to squeak, to chirp” as a verb). In some of these languages, namely Burarra, Gurindji, and Muna, ‘voice’

is in addition colexified. This is also the case, without colexification of ‘throat,’ in Katcha, Koyraboro Senni, Ngambay (which colexifies also ‘melody’), Rotokas, Aymara, and Cayapa.

Abipón, Aguaruna, Arabela, Aymara, Piro, Ancash Quechua, Tehuelche, Toba, Wayampi, and Hani colexify ‘neck’ with ‘nape’ (Xicotepec de Juárez Totonac furthermore with ‘nape of animal’ in particular), and in Cubeo the meanings are expressed using the same root suffixed with different classifiers. This pattern is also attested in Indo-European (Buck 1949: 231-232). Haida in specific contexts colexifies ‘neck’ with ‘head,’ while Rama has a term betraying this association by a complex term (*kíngkat* /*kíng-kát*/ ‘head/top-tree/foot’). Yuki and Hawaiian colexify ‘neck’ with ‘Adam’s Apple’ (alongside ‘throat,’ note also that Wappo *hóch* is glossed as “front of the neck, area of the Adam’s apple”). Rotokas *kopa iro* is analyzable as ‘swallow/ingest belt’ and colexifies ‘windpipe,’ ‘trachea,’ and ‘esophagus,’ while Tehuelche colexifies ‘neck’ and ‘to swallow’ (alongside other meanings) directly.

Basque, Bezhta, Haida, Tuscarora, San Lucas Quiaviní Zapotec, Bislama, Hawaiian, and Malagasy colexify ‘neck’ with ‘collar’ (and Greek with “neckline of garment” and Hawaiian also with “neck of a shirt or dress,” see also Buck 1949: 232 for this association in Latin and Middle High German; it may also be present in Armenian), Buli, Muna, Abzakh Adyghe, Badaga, and Basque with ‘neck of vessel’ (often in particular of a bottle), and Abzakh Adyghe and Bororo with ‘top’ in general (note also the Rama term mentioned above and that Embera *nrĩ māsí* appears to contain *nrĩ* ‘above’). The relevant Ngambay and Ngaanyatjarra terms can also assume the meaning ‘behind.’ Kyaka colexifies ‘neck’ with ‘stem,’ Kosarek Yale with “lower part of tree trunk,” and Yanomámi has *ora-hitho* ‘upper.extremity-petiole’ (compare also the formally redundant Pipil term *kech-kuyu* ‘neck-plant/stem,’ the simplex *kech* occurs only in compound); moreover, Carrier *-tšiltcen* contains *tcen* ‘stick.’ Finally, presumably accidentally, Kapingamarangi and Samoan colexify ‘neck’ with ‘rain.’

Other associations include: Efik *itōñ* is inter alia also used to refer to a “connecting part,” Hausa *wuya* also means ‘ford’ in the dialect of Sokoto, and Koyraboro Senni *jinde ~ jinda* also ‘prow of boat.’ Ngambay *gu* also means ‘chest,’ and *ngàndàra* is also an ideophone meaning “with a hoarse voice.” Buin *ou* also means ‘carry on shoulders,’ and is the name of “a form of binding for sago thatch” inter alia, while Burarra *jawa*, as a verb, means ‘to bleed, exude fluid.’ Muna *dodo-ha* is analyzable as ‘cut.off-LOC,’ the basis for the association being that the neck is the “spot where animals are cut.” Abzakh Adyghe *læcʰ* is analyzable as /l(e)-cʰ(e)/ ‘flesh-body.related,’ and *pse* also means ‘door,’ ‘to knead, kneading,’ and ‘to creep.’ Basque *lepo* can also refer to the ‘shoulder’ and ‘back.’ Laz *ali* also is used with reference to any sort of constriction, and Sora *saŋka:n* also means ‘fathom.’ Cashinahua *texu* is also used with the meaning ‘handle of a stool.’ Cavineña *erumu* also means ‘landmark’ and ‘district’ (original Spanish gloss is ‘coto’), Kaingang *nunh* also means ‘to pass over to another path,’ and Wayampi *alipi* also means ‘grape of fruit.’ Fijian *domo*, as a verb, means ‘to desire, lust after.’ Presumably by homonymy given the shortness in terms of segments, Lenakel *rou* also means “[c]hase (in order to catch).” Mandarin *jing3* also means ‘spring, well’ (due to phonological collapse of erstwhile distinct etyma, Pulleyblank 1991: 159), and Lesser Antillean Creole French *kou* also ‘to strike, hit’ among other meanings.

128. *The Nipple*

Representation: 64%

Motivated: 74.2%

Thereof Analyzable: 58.2%

Thereof Colexifying: 17.2%

Thereof by Contiguity: 25.7%

Thereof by Similarity: 26.8%

Recurrent associated meanings: breast, point/tip/end, milk, udder, eye, head, face, suck, mouth, button, nose, seed, front

Unsurprisingly, the most common association for this meaning is that with ‘breast’ (and sometimes also ‘teat,’ an association disregarded in the following). It occurs by colexification in Dongolese Nubian (where the relevant term *ért(i)* also means ‘dirt, dirty’), Swahili, Yoruba, Lavukaleve, Northeastern Tasmanian, Basque (with optional complex terms of the major types discussed below), Khalkha, Kolyma Yukaghir, Ineseño Chumash, Nuuchahnulth, Wintu, and Sedang, which colexifies also “to pound, to beat.” By far more frequent, however, are complex terms of the lexical type betraying this association. Within this class of terms, the most frequent meaning of the second element is ‘tip,’ ‘end,’ or ‘point,’ as in Kaluli *bó si /bó sí ~ tí/* ‘breast tip.’ Terms with such structure are also attested in Buli, Rendille, Baruya, Kyaka, Ket, Biloxi (dubiously), Chickasaw, Haida, Kashaya, Kiowa, Lakhota, Lesser Antillean Creole French, Carib, Guaraní (where an additional diminutive suffix is present), Jarawara, Miskito, Wichí, Lenakel, and Rotuman; a variant of this pattern is found in Kapingamarangi, which has *mada-lili* ‘end-seize.’ In addition, the association is realized formally by colexification in Hawaiian, and Rotuman colexifies ‘end’ with ‘nose,’ which is why its term also bears an association with this meaning; Ngaanyatjarra colexifies ‘nose’ and ‘nipple’ directly, with secondary associations to ‘beak’ etc. typical for terms for ‘nose.’ Terms like Yir Yoront *thayn-mel* ‘breast-eye’ are also frequent, occurring also in Badaga (where the term denotes the ‘areola’ rather than the ‘nipple’ itself), Kiliwa (where in one variant, an additional constituent meaning ‘ball’ is present), Bislama, Fijian, Samoan, Takia, and Tetun, while ‘nipple’ and ‘eye’ are colexified directly in Burarra and Hawaiian (with typical other secondary associations for ‘eye,’ compare § 6.2.3.1.). Furthermore, Jarawara has *yohari noki/yohari noko*, where *yohari* is ‘breast’ and *noki/noko* colexifies ‘eye’ and ‘seed’ inter alia. ‘Seed,’ ‘nut,’ ‘kernel,’ and ‘nipple’ are colexified in Gurindji. Burarra also has the alternative term *dombu angalginy* ‘developed nipple’ containing *dombu* ‘mudskipper’ (a type of amphibious fish with protruding eyes) and *galginy* ‘eye.’ Due to colexification with ‘eye,’ there are also associations with ‘face’ by complex terms in Toaripi, Jarawara, Fijian, Hawaiian and Samoan, and by colexification in Ngaanyatjarra.

There are further common body-part metaphors realized by complex terms with ‘breast’ acting as the contiguity anchor. Copainalá Zoque, for instance, has *tzu'tzi-pocac* ‘breast-head;’ such terms are also featured in Kosarek Yale (where ‘head’ is colexified with ‘front,’ compare for this Toaripi *omopa kō* ‘front/face breast/udder/milk’ and *kō opa* ‘breast/udder/milk front/face’), Itzaj, Yaqui, Hupda, Ancash Quechua, Toba, Malagasy, and Mandarin, which also has the variant *nai3-tou2* ‘milk-head.’ Efik colexifies ‘nipple’ with ‘mouth’ (and the term also has other secondary associations due to the meaning ‘mouth,’ cf. section 124 and § 6.2.3.2.), Buli has *biisi noai* ‘breast mouth/tip,’ Wappo *huy-nán* ‘breast-

mouth,’ and Yay *paak’ nen*<sup>3</sup> ‘mouth breast,’ and there is a semianalyzable term in Guaraní. Japanese has *chichi-kubi* ‘breast-neck.’ Other complex terms with ‘breast’ include Katcha *bibala ma thonogo* ‘child GEN breast,’ Mbum *njá-k-pâm* ‘entrance/verge-breast,’ Dadibi *bono si ame* ‘rope two woman.breast’ for ‘man’s breast nipple’ specifically, One *nimna sola* ‘breast flower,’ Sora *əjo’me:n /ə-’jo:-me:-n/* ‘POSS-unripe.fruit-breast-N.SFX,’ Cheyenne *he’e-vone* ‘female-breast,’ Kiliwa *ny+mayu=p+?uwn* ‘POSS+breast=MP+ball/knot,’ Pawnee *iit-paca* ‘breast/suck-be.a.bump,’ Hani *aqqul qulsiq*, derived from *aqqul* ‘breast, milk, sweet’ by partial reduplication and the classifier for round things *siq*, Manange *ʒɲjokro 2p<sup>h</sup>uɲ* ‘breast egg,’ Tetun *susun-lahat* ‘breast-shrimp.net’ (though note that *laho* ‘mouse, rat, muscle’ has a variant *laha*), and White Hmong *txiv-mis* ‘fruit-breast.’ There are semianalyzable terms in Guaraní and Great Andamanese, and Yanomámi has *suhe u ka* ‘breast liquid opening,’ with *suhe u* meaning ‘milk.’ This, in fact, is another natural major association for the ‘nipple,’ as already seen in terms from Kyaka, Hani, and Mandarin. ‘Nipple’ and ‘milk’ are colexified in Yoruba, Lavukaleve, Basque, Kolyma Yukaghir, Ineseño Chumash, Nez Perce, Nuuchahnulth, and Wintu, while Santiago Mexquititlan Otomí has *do-ba* ‘stone-milk’ and Central Yup’ik *emulek /emuk-lek/* ‘milk-one.having;’ the association is also present due to colexification of ‘milk’ and ‘breast’ in analyzable terms in Kyaka, Toaripi, Kosarek Yale, Kiowa, Fijian, Lenakel, Rotuman, Samoan, Tetun, and Bislama.

Further, Khoekhoe, Yoruba, Khalkha, Nez Perce, and Wintu colexify ‘nipple’ with ‘udder’ (Khoekhoe with ‘nipple of man’ specifically), and the association is realized by analyzable terms due to colexification of ‘udder’ with ‘breast’ in Kyaka, Toaripi, Lesser Antillean Creole French, Toba, and Rotuman. There may be further cases of this association, namely those in which terms for ‘nipple’ are glossed also as ‘teat’ or ‘teta’ in Spanish; these are not taken into account here because without further evidence it is not clear whether this gloss is merely a near-synonym for ‘nipple.’ Badaga and Khalkha colexify ‘nipple’ with ‘button’ alongside other meanings, among them ‘bullet’ in Khalkha. Kolyma Yukaghir *ibiši*, colexifying ‘nipple’ with ‘breast’ and ‘milk,’ is derived from a verb meaning ‘to suck’ (compare the Pawnee term just mentioned above). The association is present due to colexification with ‘breast’ in Itzaj, Lesser Antillean Creole French, and Bislama, and a semianalyzable term of this kind is found in Buin.

Other associations include: Kanuri *nzàmbòrà̀m* contains elements meaning ‘place of’ and ‘birth,’ Kosarek Yale *selkedek* is the name for a species of raspberry, and is used with the meaning ‘nipple’ in dancing songs. Badaga *ma:ru* is also a unit of measurement. Greek *rōga* also means ‘grape’ and *thilí* also ‘papilla,’ while Cheyenne *he’enénestôtse /he’e-nén-hestôtse/* is analyzable as ‘female-nurse-thing,’ Nuuchahnulth *?inma* also denotes “Nob Point, where white powder seems to run out of a breast-shaped rock,” while Oneida *ohnihsya?* has been extended to also denote ‘baby bottle.’ Santiago Mexquititlan Otomí *ts’u* also means ‘tail.’ The Wintu root *?Em* also bears the meaning “hold pectorally, carry something in the arms, embrace.” Embera *hú*, associated with different genders, can also mean ‘cavity, hollow,’ and ‘chest,’ Piro *gikatu* also means ‘stalk,’ and Wayampi colexifies ‘nipple’ with ‘stamen.’ Finally, Hawaiian *ōmaka* inter alia also means ‘building, beginning, source’ and ‘foreskin.’



129. *The Nostril*

Representation: 70%

Motivated: 78.7%

Thereof Analyzable: 76.8%

Thereof Polysemous: 2.9%

Thereof by Contiguity: 57.1%

Thereof by Similarity: 9.9%

Recurrent associated meanings: nose, hole/opening, mouth, burrow, cave, track, eye, house

Terms for 'nostril' are overwhelmingly complex, consisting of the respective terms for 'nose' (ignoring secondary meanings such as 'beak' etc. in the ensuing discussion) and 'hole' or 'opening,' as in Laz *çxindi xuntula* 'nose hole.' Terms with this structure are also found in Buli, Efik, Hausa, Kanuri, Khoekhoe, Mbum, Swahili, Yoruba, Anggor, Berik, Gurindji, Kwoma (where the relevant term also means "hole bored in nasal septum"), Lavukaleve, Mali, Muna, One, Sentani, Toaripi (where 'opening' is colexified with 'mouth,' compare also Takia *ŋdu-n awa-n* 'nose-3SG mouth-3SG; terms betraying an association with 'mouth' are also featured in Lesser Antillean Creole French, Pawnee, and Malagasy), Basque, Japanese, Ket, Welsh, Kolyma Yukaghir, Biloxi (where the meaning of the other constituent next to 'nose' is more precisely "natural orifice in the human body"), Upper Chehalis, Chickasaw, Highland Chontal, Haida, Itzaj, Kashaya, Kiliwa (for the 'inner nostril,' there is the term *phi?kw-s-?+han* 'nose=WH-INST/LONG-DN+rise.slightly' for 'outer nostril'), Kiowa, Lake Miwok, Lesser Antillean Creole French, Nez Perce, Oneida, Santiago Mexquititlan Otomí, Pawnee, Wintu, Yaqui, Bora (where segmentation is uncertain), Bororo, Carib, Guaraní, Hupda (where 'hole' is colexified with 'house,' for which compare Rama *ngú-ri* 'house-wet'), Jarawara, Miskito, Piro (where 'hole' is colexified with 'anus'), Imbabura Quechua, Toba, Tsafiki, Yanomámi, Bislama (where 'hole' is colexified with 'earth oven, cooking pit'), Fijian (where 'hole' is, as in Wintu and Sko, colexified with 'cave,' compare also Aguaruna *nuhí waa-ŋ-ti* 'nose cave-POSS-3POSS'), Hani, Hawaiian (where 'hole' is colexified with 'door' and other meanings, and the complex term can also refer to a "hole in pearl-shell shank"), Manange, Mandarin, White Hmong, Rotuman, Tetun, Vietnamese, and Yay. Semianalyzable terms containing an element meaning 'hole' are found in Upper Chehalis, Cayapa, and Wayampi, and one featuring a classifier for holes in Arabela. In Buli, Khoekhoe, Kwoma, Basque, and Toba, 'hole' is colexified with 'burrow,' and hence their complex terms for 'nostril' also betray an association with this meaning. Further variants of the pattern are Badaga *sivilu tolle* 'breath hole.for.nosering.or.earring,' and San Mateo del Mar Huave *miwirij oxing*, containing *wirij* 'hole' and *oxing* 'point.' Similarly to the situation in Piro, in Chukchi, the relevant term may be connected etymologically to 'anus,' which may itself have born the general meaning 'hole' at an earlier stage in the language's development. Baruya has *siduta* /sinna-tuta/ 'nose-track,' and a term with this structure is also attested in Kyaka. Katcha has *æ mǝ mbǝrǝ* 'eye GEN nose,' and a term with identical structure is found in Bislama; the pattern is also etymologically detectable in Tuscarora. Moreover, Swahili has *mwanzi wa pua* 'bamboo of nose,' Meyah *osúm efesi* 'nose interior,' Nivkh *vix kut* 'nose fall/come.out,' Pipil (Santo Domingo de Guzmán dialect) *-bentanah-yak* '-window-nose,' Chayahuita *nitëana*, analyzable as /nitë'-ana/ 'nose-

CLASS.AROUND,' Wichí *tonhespe*,' containing *nhes* 'nose' and a locative suffix meaning 'on,' and Great Andamanese *igchōrongal'ārjāg*, containing *igchōronga* 'nose' and *jāg* 'chink, gap.'

Ngambay, Yir Yoront and Embera colexify 'nose' and 'nostril' directly (Yir Yoront also "nose-like protrusions at the fronts of animals;" there is a redundant complex term with 'hole' for 'nostril' specifically), in Cubeo, the same root yields the meanings 'nose' and 'nostril' (and "'nose' of an airplane") depending on the classifier suffixed, and semianalyzable terms where the identifiable constituent is 'nose' are found in Bezhta, Upper Chehalis, and Cavineña.

Other associations include: Sko *loelóng* appears to contain *long* 'hole, cave' and, curiously, *loe* 'ear.' Welsh *ffroen* also means 'muzzle of gun.' Central Yup'ik *curlu* ~ *curluq* also means 'sinus' and 'head of pike fish,' and *paciguaq*, another Central Yup'ik term for the concept, is analyzable as /pacik-(ng)uaq/ 'gill-imitation.'

### 130. *The Pupil*

Representation: 48%

Motivated: 75.4%

Thereof Analyzable: 65.7%

Thereof Colexifying: 9.6%

Thereof by Contiguity: 27.4%

Thereof by Similarity: 40.1%

Recurrent associated meanings: eye, black, eyeball, child/son/daughter, seed/grain, iris, small, egg, person, spot

The 'pupil' is a meaning expressed frequently by metaphor-driven complex terms (Tagliavini 1949, Brown and Witkowski 1981), with 'eye' acting as the contiguity anchor. Of these, the association with 'child' or more specifically, 'son' or 'daughter' is most frequent, due to the small reflection of oneself in the interlocutor's eye (see Urban forthcoming for discussion of the probable cultural underpinnings). For instance, Berik has *nue tan*, presumably analyzable as /nue tane/ 'eye child/niece/nephew.' Such terms are also found in Katcha, Mbum, Kaingang, Toba, and Tetun, while Basque colexifies 'pupil' with 'child, doll,' and Greek with 'girl, daughter, virgin.' Moreover, there is a semianalyzable term in Yay, and Sedang has *kón ngái*, where *kón* is 'gibbon, spider monkey,' but note *kon* 'child.' There are also other associations with human beings, namely in Muna (*ka-mie-mie-no mata* 'DIM-person-DIM-POSS eye,' Kolya Yukaghir (*and'an-šoromə* 'eye-man'), Blackfoot (by colexification), Fijian (*yaloyalo turaga* 'reflection/image chief'), and Hawaiian (*ki'i ōnohi* 'image/doll eyeball;' *ki'i* also has some other meanings). Buin *tuutuu*, reduplicated from *tuu* meaning 'water' inter alia, is also an epithet for (plumb) children. Buli has *num-buli*, where *num* is 'eye' and *buli* can refer to a 'kid, young goat,' but also to the Buli language inter alia. Otherwise, 'seed' or 'grain' is common as the source concept, as in Copainalá Zoque *witambuj* /witam-puj/ 'eye-seed/pit.' Such terms are also found in Hausa, Kyaka (where 'seed' is colexified with 'egg' and other meanings, which is also the case in Wintu; compare Yoruba *eyinjú* /eyin-ojú/ 'egg-eye'), Ancash Quechua (where the term colexifies 'iris,' as is the case in Khoekhoe, Rendille, Baruya, Wintu, and Cubeo), Guaraní (though note that *ta'yi* 'seed' is a diminutive of *ta'y* 'son, clot'), and Tsafiki; Biloxi has *tûtçû"su'sûpi* /tûtçû"su sûpi/ 'eye seed black.' In fact, 'black' is the major non-metaphorical association for the

meaning ‘pupil.’ Complex terms such as Sora *gajar-‘mad-ən* ‘eye-black-N.SFX’ (here, the analysis is tentative; the first constituent may also be *ga’jar-* ‘to turn, to reel,’ ‘to contain a hole’), are featured in Dadibi, San Mateo del Mar Huave, Kiowa, Nez Perce, Aguaruna, Piro, Rama, and Great Andamanese, while Upper Chehalis has a complex term involving a root meaning ‘black’ and a lexical affix meaning ‘ridge, basket, trap’ (incidentally, a term with the same lexical affix also exists based on the root for ‘white’). Lenakel has *nouanhlnakanimr-in ir-apin* ‘eyeball-?? NMLZ-black/dark,’ and furthermore, Rendille has *daáyto /daáyi-to/* ‘black-thing’ and Hani *miavneev neev-nav* ‘eye/eyeball RED-black.’ There is a semianalyzable term where the identifiable constituent means ‘black’ in Khoekhoe. Another recurrent non-metaphorical association is that with ‘smallness,’ as betrayed in Kolyima Yukaghir *jukud-ənd’ə* ‘small-eye,’ also found in Highland Chontal (where the putative constituent *galninuh*, glossed as ‘little one’ more precisely is borrowed from Span. *niño* and the term hence very likely a loan translation), San Mateo del Mar Huave, Cavineña, and Samoan. There are also many unique metaphor-driven conceptualizations of the ‘pupil,’ realized by complex terms with ‘eye’ acting as the contiguity anchor in the languages of the sample. Dongolese Nubian has *míssi-ŋ-géde* ‘eye-GEN-circle/rim,’ Kwoma *miyi noku sobo* ‘eye sago.palm pale’ (*sobo* also has many other meanings), Abzakh Adyghe *nek° /ne-k°(e)/* ‘eye-middle,’ Badaga (*kan*) *maŋi* ‘(eye) gem,’ Welsh *cannwyll y llygad* ‘candle DET eye,’ Itzaj *tz’u’ ich* ‘center/heart eye,’ Kashaya *huʔu’ šihta*, /huʔuy šihta/ ‘eye bird,’ Tuscarora *yučisnúhku? ukáhrakew*, containing the roots *-čisnúhkw-* ‘comma, speck, spot,’ *-ur-* ‘to cover,’ and *-kah(r)-* ‘eye’ (for the association with ‘spot,’ compare colexification of ‘spot,’ ‘fawn,’ and ‘pupil’ in Wintu), Aguaruna *jií wincháji*, containing *jií* ‘eye’ and *winchá* ‘luminous,’ Arabela *namijia nunetejojua* ‘eye that.which.causes.to.sparkle,’ Miskito *nakra yula* ‘eye companion’ (with *yula* also having other meanings), Yanomámi *mamo ishiishi* ‘eye coal/soot,’ and Bislama *ston blong ae* ‘stone poss eye.’ There are semianalyzable terms with ‘eye’ being the meaning of the identifiable constituent in Efik, Khoekhoe, Carrier, Upper Chehalis, Highland Chontal, Cayapa, Bwe Karen, and Sedang.

‘Pupil’ is colexified with ‘eyeball,’ where relevant by a motivated term as discussed elsewhere in this paragraph, in Hausa, Katcha, Yoruba, Welsh, Wintu, and Guaraní; compare again Lenakel *nouanhlnakanimr-in ir-apin* ‘eyeball-?? NMLZ-black/dark’ as well as Hawaiian has *‘ōnohi maka* ‘eyeball eye.’

Other associations include: Khoekhoe *âutsitgares*, which also means ‘lens of eye,’ contains *âu* ‘congealed, solidified,’ and Kwoma *gwadiimay* is also the ‘generic term for spider.’ Ngaanyatjarra *tiruny(pa)* also denotes “wild onion, onion grass.” Ineseño Chumash *xutaš* is also the name of the fruit of the coffeeberry (as well as ‘evening star, Venus’ and “Earth mother”); similarly, Haida *hldaandaraay* contains *hldaan* ‘blue huckleberry.’ Central Yup’ik *takvik ~ takviun* is analyzable as /taku-vi-(u)n/ ‘check.fishtrap.or.fishnet-??-device.for.’ Hawaiian *ki’i ‘ōnohi* can also be used to refer to a ‘beloved person.’ Rotuman *rito* also denotes the “young leaves (of coconut or other palm) just as they are coming out at the top of the trunk (in the centre).”

131. *The Pus*

Representation: 79%

Motivated: 23.4%

Thereof Analyzable: 5.6%

Thereof Colexifying: 17.8%

Thereof by Contiguity: 6.9%

Thereof by Similarity: 9.5%

Recurrent associated meanings: infected/infection, rot/rotten, wound/sore,  
semen, dirt/dirty, boil, phlegm, resin, snot, brain, blood, water/liquid

Cheyenne, Comanche, San Lucas Quiaviní Zapotec, Ancash Quechua, Wayampi, and Yanomámi colexify, by provenience contiguity, ‘pus’ with ‘infect, (be) infected, infection;’ Yanomámi also has the complex term *niyo-niyo u-pě* ‘infect-RED liquid-PL.HETEROGENOUS.’ Alongside Santiago Mexquititlan Otomí, which directly colexifies ‘pus’ and ‘wound,’ there are also several languages which have complex terms for ‘pus’ with ‘wound’ or ‘sore’ acting as contiguity anchor. Highland Chontal has *lija ga-wi* /lijabíh la-wi/ ‘lime.water.from.cooking.corn wound,’ San Mateo del Mar Huave *aonts najloc ~ aonts necoy* ‘excrete wound’ (note also that there is a semianalyzable term with a constituent meaning ‘excrement’ in Sáliba and that this association is etymologically recoverable in Dongolese Nubian), Kashaya *maʔsa tʰo*, analyzable as /maʔsa tʰoʔ/ ‘sore rot,’ and Tetun *kanek-been* ‘wound-liquid.’ Moreover, there is a semianalyzable term where the identifiable constituent means ‘sore, wound’ in One. Alternatively, Tetun also has *been-tasak* ‘liquid-septic,’ and an association with ‘water’ is also found in Sora: *gurda:n* ‘pus, juice, sap’ contains *d’a*- ‘water’ and either *gʷu*- ‘to ripen’ or *gu*-, a variant of *gor*- ‘to cut vertically soft things’ (for this term, note that Mali colexifies ‘pus’ with “white sap from the breadfruit tree” in particular). There is a semianalyzable term with ‘water’ as the meaning of the identifiable constituent in Sko, and one in Abzakh Adyghe which colexifies ‘pus’ with ‘lymph.’ The association with ‘rot, (something) rotten’ in Kashaya is not unique: Gurindji, Ngaanyatjarra, Santiago Mexquititlan Otomí, and Tuscarora colexify these meanings, and Tetun has *raan-kroek* ‘blood-rotten’ alongside *raan-mutin* ‘blood-white.’ This, in turn, is not the only association with ‘blood,’ which is colexified with ‘pus’ in Kosarek Yale.

There are also other associations with body fluids and soft body parts, which are more frequent cross-linguistically. Baruya and Manange colexify ‘pus’ with ‘snot’ (this may also be the case in Cayapa, where the terms are however not quite identical), Ket, Lake Miwok, Carib, and Yay with ‘semen’ (Yay with ‘pus running from the flesh’ specifically), and Ket and Lake Miwok colexify ‘pus’ with ‘brain.’ Badaga and Wintu colexify ‘pus’ with ‘phlegm.’

Furthermore, Sahu and Abzakh Adyghe colexify ‘pus’ with ‘boil,’ Hawaiian has *pala-hēhē* ‘yellow-boil’ (*pala* can also assume the meaning ‘rotten’ when speaking about taro corms inter alia), and there is a semianalyzable term containing an element meaning ‘abscess’ in Kaingang. Ngaanyatjarra, Rama, and Bislama colexify ‘pus’ with ‘dirt(y)’ (in Swahili, the term for ‘pus’ is a loanword from Arabic originally meaning ‘dirt, filth’). Piro colexifies ‘pus’ with ‘resin’ and ‘rainbow,’ and the former association is also present in Sora.

Other associations include: Bakueri *ijá* also denotes “a bunch of oil-palm nuts.” Hausa *’diwa* is also the name of a red grass species, and dialectally assumes the meaning ‘anus,’ while *mugunya* is also the feminine form of *mugu* ‘evil.’ Muna *taghirao* ‘pus from acne’ also means ‘soft meat inside an old coconut’ (compare *taghi* ‘belly, stomach’?), Ket *do’ŋ* also means ‘three,’ Khalkha *idege(n)* also “food, nourishment, provisions,” ‘kernel of nut,’ and ‘tannin,’ and for Khalkha *øgeri* and *øgesyn* compare *ø ~ øge* ‘fault, roughness, unevenness.’ Sora *tule:dan* colexifies ‘pus’ with ‘gum,’ and Pipil *te:mal* is derived from a verb meaning ‘fill.’ The Tuscarora root *-atkehθr-* also yields the meaning ‘leather wood.’ Central Yup’ik *imaq* means ‘content’ and also ‘bullet’ and ‘ocean.’ Wayampi appears to colexify ‘pus’ with ‘street’ inter alia, Great Andamanese *mûn* also denotes ‘marrow,’ Fijian *nana* is also an ‘affectionate word for mother,’ while Hani *biaol* also means ‘to fly.’ Bwe Karen *mi* also means “be full of pus” and ‘be ripe’ inter alia, Sedang *hẽ* also ‘saliva,’ and Lesser Antillean Creole French *matie* also ‘matter, material’ and ‘topic.’

### 132. *The Rib*

Representation: 90%

Motivated: 33.6%

Thereof Analyzable: 23.1%      Thereof Polysemous: 12.0%

Thereof by Contiguity: 27.4%      Thereof by Similarity: 4.5%

Recurrent associated meanings: side/flank, bone, chest/thorax/ribcage, fence

By far the most common structure in motivated terms for ‘rib’ in the languages of the sample is for them to be analyzable of the lexical type, with constituents meaning ‘bone’ and ‘side’ or ‘flank,’ as for instance in Kildin Saami *jërrh’t-tāxx’t* ‘flank-bone’ (both are common associations in Indo-European, Buck 1949: 208). Such terms are also featured in Efik, Kanuri, Koyraboro Senni, Mbum, Yoruba, Sahu, Miskito (which also has the alternative term *tnaya mina* ‘side dent’), Tsafiki, Wichí, Hawaiian (where the relevant term also means ‘spareribs’ and ‘wife,’ because of the biblical motive of Eve having been created from Adam’s ribs), Malagasy, Rotuman, Takia, Tetun, and Yay. There are also, alternatively, some languages with complex terms involving ‘bone,’ but not ‘side.’ These are One (*nenki amna*, with *amna* meaning ‘bone’ and *nenki* also occurring in *moru nenki* ‘vertical thatch braces,’ similarly, Ineseño Chumash colexifies ‘rib’ with “verticals of house frame,” Angkor *ngeremb* is also glossed as “frame of mbisu” and *kwansatha* in the Piro term for ‘rib’ *kwansathapu* means also ‘ribs of roof;’ this association is paralleled in Indo-European, Buck 1949: 208), Badaga (*nenjilu* /*nenju-ilu*/ ‘chest-bone’), Ket (*ulat* /*ul-a’d*/ ‘straight-bone’), Kolyma Yukaghir (*nugod’id-amun* ‘thigh-bone,’ colexifying ‘thigh’), Kapingamarangi (*iwi di wogowogo* ‘bone ART ribcage’), Manange, which has (*ŋoytse*) *1krẽ 2nokrẽ* ‘(under.waist) hips bone,’ and Samoan (*ivi ‘aso’aso* ‘bone variety.of.yam’). Semianalyzable terms with ‘bone’ as the identifiable constituent are attested in Efik, Sora, the Cuisnahuat dialect of Pipil, Copainalá Zoque, and Guaraní.

Conversely, there are also languages where the relevant terms betray an association with ‘side,’ but not with ‘bone.’ Kyaka, Ngaanyatjarra, Sko, Basque, Greek, Khalkha, Nivkh, Bororo, Bislama (by a rare term), and Hani directly colexify ‘rib’ with ‘side (of

body)’ (Basque also with ‘hulk’ and ‘point of view, viewpoint’ and Khalkha also with ‘wing’ and “spur of a mountain between two valleys”), Chukchi has *γəto-lqəl* ‘side-material.for,’ Piro *kwansatha-pu* ‘side-shape.of.bean.or.banana,’ and there is a semianalyzable term in Gurindji.

Moreover, Buli, Ngambay, Buin, Rotokas, Kiliwa, Xicotepec de Juárez Totonac, Bororo (by a semianalyzable term containing the constituent ‘tree’ and colexifying ‘trunk’), Fijian, and Mandarin colexify ‘rib’ with ‘chest’ and/or ‘thorax, ribcage’ (Ngambay also with ‘basket’ and ‘measles’), while in Cubeo, the same root yields both meanings depending on the suffixed classifier, Bora has *mijowa*, perhaps analyzable as /mijco-gwa/ ‘chest-scm.2d.straight,’ and the association is also present in the Badaga and Kapingamarangi terms mentioned above. Finally, Lavukaleve colexifies ‘rib’ with ‘fence,’ and analogously, Muna has *karakara*, with the apparent reduplication base *kara* meaning ‘yard, yard fence.’

Other associations include: Hausa *hak’ark’ari* also means ‘pneumonia,’ Kaluli *sidif ~ tidif* appears to contain *sí ~ tí* ‘tip,’ and Meyah *osrój* also means ‘fiancée, boyfriend, girlfriend.’ Yir Yoront *pawrrmel* contains *mel* ‘eye,’ while one constituent of the Abzakh Adyghe term *cağe* is *ce* ‘fiber, blade.’ Khalkha *xabirya(n)*, apparently derived from *xabir- ~ xabira-* ‘to whet, grind, rub, touch lightly in passing, for animals to stand close to one another,’ also means ‘wing,’ *xabisu(n)*, which seems to be derived from *xabi* ‘vicinity, neighborhood,’ also means ‘womb, uterus,’ and Welsh *asen* also ‘she-ass.’ Central Yup’ik (Yukon, Lake Iliamna and Nunivak Island dialects) *inarun* is analyzable as /inar-te-(u)n/ ‘lying.down.act.on.one.so.as.to.cause.it-device.for.’ Arabela *riuquiocua ~ riquiocua* is analyzable as /riuquionu-cua/ ‘bend.something.to.form.a.receptable-CLASS.MOULD.’ Cayapa *vi’chi* might consist of *vi* ‘chaquira’ and *chi* ‘tree, wood,’ and Rama *palkát* contains *kát* ‘tree, stick.’ Fijian *sarisari* also denotes the “ribs” of a boat, while *waqawaqa*, denoting both ‘ribcage’ and ‘rib,’ is reduplicated from *waqa*, meaning ‘box, container’ inter alia. Bwe Karen *-we* is glossed inter alia as ‘shore, bank, waterside’ in the Bwe Karen-English section of the consulted source, while Lenakel *nakau* also means ‘midrib of coconut’ and ‘side of mountain.’

### 133. *The Saliva*

Representation: 82%

Motivated: 34.3%

Thereof Analyzable: 18.0%

Thereof Colexifying: 16.4%

Thereof by Contiguity: 11.3%

Thereof by Similarity: 16.2%

Recurrent associated meanings: water/liquid/juice, phlegm/sputum, mouth, foam, spit, synovial, soup

Terms for ‘saliva’ (‘spittle,’ ‘drool,’ ‘slaver’) of the analyzable lexical type in which one constituent means ‘mouth’ (possibly with associated extensions, cf section 124 and § 6.2.3.2.) and the other ‘water,’ ‘juice,’ or ‘liquid’ more generally are relatively frequent. For instance, Yuki has *nan-uk* ‘mouth-water,’ and such terms are further found in Koyraboro Senni, Ngambay, Kyaka, Abzakh Adyghe, Upper Chehalis, Chickasaw, Itzaj, Kiliwa, Wappo, Jarawara, Maxakalí, Miskito, and Lenakel. San Mateo del Mar Huave has *aonts ombeayaran*

'excrete mouth,' and semianalyzable words with the identifiable constituent meaning 'mouth' are attested in Nunggubuyu and Cashinahua. Furthermore, Kosarek Yale *sulu-mak* appears to be analyzable as 'strong-water,' Chayahuita has *iro-i* 'cough/flu-CLASS.LIQUID,' Tsafiki *pi'pi* appears to be reduplicated from *pi* 'water, liquid, juice,' Yanomámi has *kahu u* 'cover liquid,' Hawaiian *wale wai* 'slime/phlegm water,' and Bwe Karen *bəyà-chí* 'person-water.' In Khoekhoe, a term for 'saliva of animals, dribble' dialectally also has the meaning 'water,' and likewise, one for 'foam, saliva' means 'water' in the Northern dialect of Ngaanyatjarra. There is a derived term in Great Andamanese, and semianalyzable terms are found in Kosarek Yale, Ket, Bororo, Piro, Rama (the relevant constituent meaning 'wet' rather than 'water'), and Tetun ('juice' is colexified with 'water' or 'liquid' in a few of the above mentioned languages, and a semianalyzable term with a constituent meaning 'juice' specifically is found in Wayampi). In Itzaj, *k'a* in *k'a' chi* 'saliva' colexifies 'liquid' with 'juice' and 'soup,' while Khalkha and *sily(su)* seems to be derived from *sily(n)* 'soup, buillon, broth' (Khalkha has another term, *nilbusu(n)* ~ *nilmusu(n)*, which also can refer to 'tears').

Otherwise, two languages of the sample, Sahu and Ineseño Chumash, colexify nominal 'saliva' with verbal 'to spit,' while in Khalkha and Fijian, there are terms for 'saliva' standing in a derivational relationship to a verb meaning 'to spit;' further, in Koyraboro Senni, Biloxi, Nuuchahnulth, and Tehuelche, there is an apparent relationship of this kind, but the precise mechanism of word formation cannot be identified (anymore). Seventeen sampled languages, Khoekhoe, Ngambay, Berik, Badaga, Khalkha, Haida, Quileute, Aguaruna, Arabela, Cashinahua, Cavineña, Cubeo, Guaraní, Huambisa, Ancash Quechua, Hani (the relevant term *zaoqpeiv* contains *peiv*, meaning inter alia 'to have a continued sickness, to have a relapse of sickness' and also functioning as a classifier for dew-drops; *zaoq* also has other apparently unrelated meanings), and Hawaiian, colexify 'saliva' and 'phlegm,' and in Embera, the same root associated with different genders yields both relevant meanings. Efik and Khoekhoe colexify 'saliva' with 'synovia.' Furthermore, there is a pattern of colexification, namely that with 'foam,' that is particularly frequent in Oceania. It is attested in Buin (which also colexifies 'lungs'), Kwoma, Lavukaleve, Ngaanyatjarra (which colexifies also 'soap powder'), Tasmanian (Western and Southeastern), Badaga, Lesser Antillean Creole French, Tsafiki, and Bislama.

Other associations include: Efik *u'döt*, derived from *dör'ö* 'be viscous, glutinous,' is a general term for "any viscid animal secretion," while Hausa colexifies 'saliva' with 'day, today.' Koyraboro Senni *hatta* also means to "gather spittle into throat before spitting" as well as 'to clear one's throat' and 'to miss target.' Ngambay colexifies 'saliva' with 'vomit,' and Yoruba *itọ* also means 'small creek.' Nunggubuyu *n<sup>a</sup>al* is a term for 'slimy substance' generally, which can for instance also refer to decaying jellyfish on the beach, Rotokas *vevega* also means 'slime' and 'semen,' Sentani *ki* also 'deed, act' and 'infant,' and Yir Yoront colexifies "vapour spouted by sea-mammal, spray spouted by sea-mammal." Badaga *ecca* ~ *eccalu* ~ *enjala* also means "impurity from contact with mouth" and 'leftover food,' and Basque *listu* also 'to fray, unravel.' The Blackfoot term *sóopoyoyihkaan* contains *soopoyóoyihkaa* 'to drool,' Kiowa *toutk'yhdl* contains *k'yhdl* 'to be wet,' and Embera *ídubá* contains *í* 'lip' and *đu* 'deep.' Guaraní *tendy* also means 'flame, light,' the Huambisa term

*saawin* contains *saawi* ‘clear, transparent,’ while Kaingang *jã-ra* appears to be analyzable as ‘tooth-jaw.’ Lengua *emanang* might contain *ema* ‘blood,’ and Great Andamanese *âkàraij* might contain *raij* ‘to shoot.’ Hawaiian ‘*ae*’ also means “sap wrung from seaweeds or leaves of plants such as taro” inter alia, and *hā‘ae* is also the name of “a variety of sweet potato from which bear [sic] was made.” Finally, Sedang *hẽ* also means ‘pus.’

#### 134. *The Scar*

Representation: 78%

Motivated: 33.2%

Thereof Analyzable: 12.3%

Colexifying: 21.7%

Thereof by Contiguity: 14.7%

Thereof by Similarity: 9.6%

Recurrent associated meanings: wound/sore, mark, cut, trace, sign, heal, footprint, scab, welt, birthmark, spot, place, lump/lumped, depression, scale, stain, shadow, old

Koyraboro Senni, Yoruba, Ngaanyatjarra, Nunggubuyu, Yir Yoront, Piro, Tehuelche, and Manange colexify, by provenience contiguity, ‘scar’ with ‘wound’ or ‘sore’ (Yir Yoront also with ‘bite of insect’). Alternatively, Sko has *píúe-lúe*, presumably analyzable as ‘wound-chop/cut.small.things,’ Kosarek Yale *meneng kikle* ‘wound cautious,’ Basque *zauri-marka* ‘wound-mark,’ Japanese *kizu-ato* ‘wound-trace’ (this is quite a frequent colexification cross-linguistically: ‘scar’ and ‘trace’ are colexified in Efik, Yoruba, Muna, Abzakh Adyghe, Toba, and Hawaiian, and Lenakel has a semianalyzable term), Kildin Saami *avv’-sajj* ‘wound-place,’ Hupda *həm bǝg* ‘wound old’ (compare also Wayampi *(l-)ɛna-ɛ* ‘(DEPENDANCE-)place-PAST’), Bislama *mak blong soa* ‘mark/spot POSS sore’ and *trak blong soa* ‘track POSS sore,’ and Tetun, like Kildin Saami, *kanek-fatin* ‘wound-place.’ Semianalyzable terms with ‘wound’ or ‘sore’ are found in Ngambay, Berik, Kaluli, One, Piro, Sáliba, and Hani, in which latter there are other semianalyzable terms involving a constituent meaning ‘be hurt, ache, be ill.’ There are also a number of sampled languages in which ‘scar’ bears an association with the meaning ‘to cut, cut.’ This is found by colexification in Koyraboro Senni (Gao dialect), Hawaiian, and Samoan, Sko has, as mentioned above, *píúe-lúe*, ‘wound-chop/cut.small.things,’ Itzaj has *kuuch xot'al tu'pal* ‘place cut/chopped extinguished,’ Santiago Mexquititlan Otomí a derived term, Aguaruna *tsupí-k-mau wakaní* ‘cut-ASP-NMLZ shadow’ (note also the colexification of ‘scar’ and ‘shadow,’ alongside ‘companion, spouse, friend’ in Cashinahua), and Kapingamarangi *di lohongo me ne wele/tuu* ‘DET position thing PERF burned/cut.’ Baruya, Nuuchahnulth, and Bwe Karen colexify ‘scar’ with ‘heal(ed),’ Ket has *binaxols* containing *bin* ‘self’ and *qol* ‘heal,’ and Fijian *i macamaca* ‘DERIV dry/healed.’ Buli, Efik, Hausa, Yoruba, Yir Yoront, Basque, Greek, Khalkha, Toba, and Lesser Antillean Creole French colexify ‘scar’ with ‘mark’ (Yoruba also with ‘impression;’ similarly, Samoan has *mā-‘ila* ‘COMPLETE-mark’), Yoruba, Greek, Highland Chontal, Lesser Antillean Creole French, and Yanomámi with ‘sign,’ Buli, Efik, Guaraní (among other meanings), and Toba colexify ‘scar’ with ‘footprint’ (furthermore, there is a semianalyzable term with a constituent with this meaning in Lenakel), Hausa and Basque with ‘spot,’ Tuscarora, Aguaruna and Hawaiian with ‘scab’ (Aguaruna also with ‘sterile,’ and Hawaiian also with ‘button, badge’ and ‘blunt, dull’ as well as ‘fish’), Upper Chehalis and Guaraní with ‘lump(ed),’



Chayahuita and Hawaiian with ‘depression,’ Muna and Tuscarora with ‘scale,’ Khoekhoe, Noni, and Greek with ‘birthmark,’ Yoruba and Hawaiian with ‘stain,’ and Muna, Nunggubuyu, and Hawaiian with ‘welt.’ Finally Yanomámi has a semianalyzable term for ‘scar’ with the identifiable constituent meaning ‘skin,’ also meaning ‘cover’ more generally.

Other associations include: Efik colexifies ‘scar’ inter alia with ‘seam,’ ‘crack,’ ‘flaw,’ and ‘chink,’ and *inia* is also the name of “a disease covering the body with large sores.” Hausa *adabali* also denotes a “patch riveted on to a damaged sword,” and *kufai* also an “old site of house, compound, or even a town” inter alia. *Tabo*, another Hausa term, is also used to refer to a favor not paid back and an old fault that can be brought up against someone. Burarra *munjakarn* contains *munjak* “hard ground on the floodplains which have dried out,” and Kyaka *mumbwua* colexifies ‘scar’ with ‘callus.’ Muna *bhili* also denotes “hanging roots, supporting roots,” and *ghana* also means ‘to not finish, not use up.’ *Pinda* is another Muna term with apparently unrelated additional meanings, among them “traditional plate made of bone.” Relevant Tasmanian terms in all varieties with the exception of the Northern (for which data are lacking) also mean ‘wart,’ ‘wrinkle,’ and ‘tail,’ and in Yir Yoront there is a semianalyzable term where the identifiable constituent means ‘mouth.’ Basque *orbain* also denotes a ‘pock.’ Greek *sīmádi* also means ‘target,’ as well as ‘omen, sign.’ Sora *dul’dulən* appears to be analyzable as /döl-döl-ən/ ‘RED-finish/bite-N.SFX,’ and *gag’garən* ~ *gal’galən* as /gag-gar-ən/ ‘RED-pierce/boar.a.hole-N.SFX;’ this term also denotes the “pits of small-pox” specifically. *Tar-pu:-n* is analyzable as ‘white-stab-N.SFX,’ and *pənuɡo:n*, yet another Sora term, is related to *’pugo-* “to be coloured, to be scarred, to be spotted, to be overcast with dust.” Blackfoot *isttsikakkssin* is derived from *isttsikakkí-* ‘to form shiny scar tissue,’ which in turn seems to contain *isttsi* ‘pain, ache, hurt’ and perhaps *wakkii* ‘to heal.’ Wintu *tʰaw* also means “grief, mourn, sad.” Central Yup’ik *qelengllak* also means ‘wrinkle’ and ‘kink,’ and *iiraq* also ‘parotid gland,’ ‘tonsil,’ and ‘side of neck.’ Copainalá Zoque colexifies ‘scar’ with ‘crab species,’ whereas Cubeo *toa-churi* is analyzable as ‘fire-CLASS.SCAR.OR.WOUND-LIKE.OBJECT.’ Piro *ptšə* also means ‘dirtiness.’ Yanomámi *kano si* also means ‘rubbish’ and “abundant fruits which one collects several times from one place,” and *tusi* ~ *tusitusi* also means “covered anew with vegetation.” Bwe Karen *gəθrətha* contains *tha* ‘up, rising.’ Figuratively, Hawaiian *ālina* also means ‘low, disgraced, dishonored,’ and *lina* also means ‘soft’ and ‘sticky’ (there are also other apparently unrelated meanings). Malagasy *hòlatra* also means ‘mushroom,’ Rotuman *pətu* also “to hit (a person’s head) with the back of the fingernail,” while Yay *pan<sup>4</sup> piaw<sup>3</sup>* is analyzable as ‘become to.sear.’ Lesser Antillean Creole French *mak* also means ‘stamp’ and ‘mole’ inter alia.

### 135. The Skin

Representation: 97%

Motivated: 73.4%

Thereof Analyzable: 3.1%

Thereof Colexifying: 70.0%

Thereof by Contiguity: 20.0%

Thereof by Similarity: 26.0%

Recurrent associated meanings: leather/hide, bark, rind/peel, shell, surface/covering, husk/chaff, scale, body, crust, clothing/cloth, complexion, pod, fingernail, scab, flesh, race, kin/family, rubber tire

Many sampled languages colexify 'skin' with other types of outer covering of living things, often with colexification of more than one additional sense. Many languages colexify 'skin' with 'bark' (and there are many other languages in which 'bark' is expressed by a complex term on the basis of 'skin,' compare section 3; this association is only weakly attested in Indo-European, namely in Germanic, according to Buck 1949: 201). The relevant languages are Bakueri, Efik, Yoruba, Anggor, Buin, Burarra, Gurindji, Kwoma, Kyaka, Lavukaleve (colexifying 'bark of coconut tree' specifically), Mali, Muna, Ngaanyatjarra, Nunggubuyu, Sahu, Sko, Sentani (alongside a particular palmtree in its entirety), Southeastern and Western Tasmanian, Toaripi, Kosarek Yale, Basque, Bezhta, Biloxi, Chickasaw, Ineseño Chumash, Comanche, Haida, Pipil, Xicotepec de Juárez Totonac (colexifying 'skin of a person's leg' specifically), Yaqui, Copainalá Zoque, Abipón, Bororo, Carib, Cashinahua, Cayapa, Guaraní, Hupda, Jarawara, Kaingang, Maxakalí, Piro, Ancash Quechua, Tsafiki, Wayampi, Yanomámi, Bislama, Fijian, Hawaiian, Bwe Karen, Lenakel, Malagasy, White Hmong, Rotuman, and Samoan. Moreover, Cahuilla colexifies 'bark' with 'skin of animals.'

Bakueri, Yoruba, Baruya, Kyaka, Mali, Muna, Ngaanyatjarra, Basque, Khalkha, Nivkh, Welsh, Chickasaw, Haida, Lake Miwok, Lakhota, Lesser Antillean Creole French, Oneida, Pawnee, Pipil, Wintu, Central Yup'ik, San Lucas Quiaviní Zapotec, Bororo, Embera, Guaraní, Jarawara, Kaingang, Rama, Yanomámi, Bislama, Fijian, Hawaiian, Rotuman, and Tetun colexify 'skin' with 'rind' and/or 'peel,' and Efik, Yoruba, Muna, Nivkh, Cahuilla, Chickasaw, Ineseño Chumash, Lake Miwok, Lakhota, Pawnee, Pipil, Xicotepec de Juárez Totonac, Yuki, Copainalá Zoque, Bororo, Cashinahua, Cayapa, Embera, Guaraní, Hupda, Kaingang, Maxakalí, Ancash Quechua, Sáliba, Tsafiki, Yanomámi, and Bwe Karen colexify 'skin' with 'shell' of various objects, e.g. a nut or crustacean. Bakueri, Efik, Yoruba, Baruya, Muna, Rotokas, Sahu, Khalkha, Nivkh, Embera, and Bislama colexify 'skin' with 'husk' and/or 'chaff' (see also Buck 1949: 201 for this association in diachrony in Indo-European), Efik, Nivkh, Biloxi, Lake Miwok, Wintu, Guaraní, Jarawara, and Bwe Karen with 'scale' (see also Buck 1949: 201 for this association in diachrony in Celtic), Kyaka, Muna, Basque, and Rotuman with 'crust,' Efik, Yoruba, and Nivkh with 'pod,' Efik and Central Yup'ik with 'scab,' Anggor, Sko, Kiowa, and Yanomámi with 'clothing' or 'cloth,' and Highland Chontal and Aymara with 'flesh' (Highland Chontal also with 'pulp').

Still more generally, terms for 'skin' are extended to 'surface' or 'covering' in general in Kwoma, Sko, Abzakh Adyghe, Basque, Nivkh, Comanche, Haida (also with 'surface of waves'), Lake Miwok, Lakhota (colexifying also 'envelope' and 'wrapping' specifically), Nuuchahnulth, Cubeo, Guaraní, Jarawara, Lengua, Maxakalí, Piro, Wayampi, Yanomámi, Bislama, Hawaiian, Kapingamarangi, Bwe Karen, and Lenakel (and 'cover' is the ultimate source of many Indo-European terms for 'skin,' Buck 1949: 200-201).

Noni, Baruya, Buin, Kwoma, Mali, Aymara, and Toba colexify 'skin' with 'body,' while Dadibi *tigi wali* is analyzable as 'body wind.' Ineseño Chumash, Cubeo, and Hawaiian colexify 'skin' with 'complexion' (as did Ancient Greek, Buck 1949: 200), and Ngaanyatjarra

and Khalkha with 'race.' By metaphorical extension, Miskito and Samoan colexify 'skin' with 'kin, family.' Oneida colexifies 'skin' with 'rubber' (by a term probably containing an element meaning 'piece of cloth, rag'), and similarly, San Lucas Quiaviní Zapotec *gui'ihdy ~ gyi'ihdy* also denotes 'plastic,' while Rotuman and Samoan colexify 'skin' with '(rubber) tire'

Another major association is, by provenience contiguity, that with 'leather' and/or 'hide' and sometimes also 'pelt' (see also Buck 1949: 200 for ample Indo-European evidence). This is found exclusively by colexification in the languages of the sample and is attested in Buli, Efik, Khoekhoe, Koyraboro Senni, Noni, Dongolese Nubian, Swahili, Yoruba, Kyaka, Muna, Sentani, Abzakh Adyghe, Badaga, Basque, Chukchi, Greek, Khalkha, Sora, Blackfoot, Upper Chehalis, Haida, San Mateo del Mar Huave, Kiowa, Lesser Antillean Creole French, Nuuchahnulth, Santiago Mexquititlan Otomí (by the term *xifni*, analyzable as /xi-fani/ 'body.hair-horse'), Pawnee, Pipil, Quileute, Xicotepec de Juárez Totonac, Tuscarora, Wappo, Yana, Yaqui, Central Yup'ik, San Lucas Quiaviní Zapotec, Copainalá Zoque, Abipón, Aguaruna, Arabela, Bora, Bororo, Carib, Cashinahua, Cavineña, Chayahuita, Embera, Guaraní, Huambisa, Jarawara, Kaingang, Lengua, Miskito, Piro, Ancash and Imbabura Quechua, Sáliba, Tehuelche (colexifying 'ostrich leather' more specifically), Toba, Tsafiki, Yanomámi, Hawaiian, Bwe Karen, Malagasy, Manange, White Hmong, Rotuman, Samoan, Sedang, and Tetun.

Other associations include: Buli *gbain* also means 'drumhead' and *gbang* also means 'paper' and denotes also several objects made of paper; it can also refer to 'gambling.' Efik *ik'pa* inter alia also means 'whip,' 'stripe,' and 'surface,' e.g. of water lying still. Hausa *fata* also means 'to wish well, well-wishing' and "blow the nose and project mucus into or upon something." Ngambay colexifies 'skin' and 'shoe,' Burarra *-maliyarra* 'skin' and "wrapper, outer case" (this term may be related to *mala* 'clan'), Muna *kuli* also means to "have a simple meal without side dish," Ngaanyatjarra *miri* also to "get a shock, be shocked" and *pangki* (Northern dialect) also 'orange.' One *tapi* also means 'thin,' and Meyah *ofos* is also used with the meaning 'mountain top.' Waris *tóvól ~ nihtóv* also means 'blood,' Abzakh Adyghe *ś'e* also 'appearance, color' and *śhe* inter alia also 'head, top,' Badaga *to:lu*, similarly to One, also 'skinny,' Basque *azal* also 'sheet,' 'case,' as well as 'rascal,' Nivkh *hatx* also 'envelope,' 'film,' and 'fur,' and Biloxi *ahi' ~ ahe' ~ ahě' ~ he* also means 'fingernails, toenails' (Rama also colexifies 'nail'), 'horn,' and 'hoof.' Kiowa colexifies 'membrane' and 'cloth, mat,' Wintu *čop* also means 'acorn' and *la'* also 'tendon' and 'string, stretch.' Copainalá Zoque *naca* is also used with the meaning 'sole.' Hupda *b'ók* also means 'dish, plate, food,' and Maxakalí *xax* also to "seek, hunt, long for." Miskito *taya* also means 'feather,' Piro *mta* also 'mat,' and Rama *uk* also 'coat' and 'nail.' Toba *l'oc* may colexify 'skin, body' with 'cloud,' though *l-* in the body-part terms is a 3<sup>rd</sup> person singular possessive prefix, and it is unclear whether the similarity with *l'oc* 'cloud,' with the consonantal onset apparently belonging to the root, is merely accidental. Wayampi *pi* also means 'to finish' inter alia. Bwe Karen *(-)kó* can also mean "to swell as the result of the presence of pus or fluid under the skin" and 'mountain,' and *phe* also "to scratch, claw, maul." Hawaiian *'ili* also means 'area, land section' and *'alu'alu*, formally reduplicated from *'alu* meaning inter alia 'sag, flabby' and 'depression, gully,' has many meanings, among them 'foetus,' while Lenakel

*nivig-* may also refer to the ‘cover of a book.’ Samoan *pa’u* also means ‘belt of machine’ and ‘foreskin,’ and White Hmong *tawv* also means ‘hard.’

### 136. *The Snot*

Representation: 75%

Motivated: 37.4%

Thereof Analyzable: 15.8%

Thereof Colexifying: 21.6%

Thereof by Contiguity: 15.0%

Thereof by Similarity: 14.3%

Recurrent associated meanings: nose, cold/flu, phlegm/mucus, blow nose, runny nose, water, pus, faeces

Two sampled languages, San Mateo del Mar Huave and Kashaya, have complex terms of the lexical type for this meaning with the constituents meaning ‘nose’ and ‘excrete, faeces.’ Furthermore, there are semianalyzable terms where the identifiable constituent is ‘faeces’ in Cayapa and Lenakel. The Cayapa term just mentioned colexifies ‘snot’ with ‘flu’ (its shape is *quijpe*, compare also *quijpe* ‘pus?’), and this or colexification with ‘cold’ or ‘have a cold’ is also found in Hausa (dialectally), Buin, Ngaanyatjarra, Yir Yoront (by a semianalyzable term also containing ‘nose’), Wintu, Arabela, Bora, Embera, Hupda, and Miskito. Sahu has *ma si’dangutu* ‘POSS have.a.cold,’ and Chayahuita *iro nitën quëran pipirinsó*, containing *iro* ‘cough, flu,’ *nitë* ‘nose,’ and *pipirin* ‘to come out.’ There is a semianalyzable term containing an element meaning ‘flu’ in Piro. In four sampled languages, Abzakh Adyghe, Itzaj (where the term means ‘clear mucus’ specifically), Kiowa, and Tetun, ‘water’ or ‘liquid’ rather than ‘faeces’ is the meaning of the second constituent, for instance Kiowa *mā-t’əy* ‘nose-water’ (a similar situation is etymologically recoverable in Tuscarora). A semianalyzable term with ‘water’ as the identifiable constituent is found in Mbum. Alongside the term with this structure which is used for ‘watery snot,’ Abzakh Adyghe also has *pe-šan* ‘nose-pus’ for ‘slimy snot,’ and ‘snot’ and ‘pus’ are colexified in Baruya and Manange. ‘Snot’ and ‘phlegm, mucus’ are colexified in Hausa, Ngaanyatjarra, Badaga, Basque (also with ‘gum, resin,’ and ‘wick’), Itzaj, Pawnee, Tuscarora, Embera, Tehuelche, and Fijian.

Upper Chehalis has *s-táp=qs* ‘CONTINUATIVE-thick(of liquid)=nose/point,’ Kiliwa *phi?=chiilq* ‘nose=boil/eruption/pustule,’ Bororo *eno bori* ‘nose wax,’ Carib *enata aikulu*, containing *enata* ‘nose’ and *aiku* ‘juice,’ Embera *kú-mór* ‘nose-INTERNAL.SUBSTANCES.OR.ORGANS.OF.BODY,’ Rama *táik síri* ‘nose slime,’ Yanomámi *hushihushihi* (compare *hushi* ‘nose’), Bislama *bata nus* ‘butter nose,’ and Samoan *isu-pē* ‘nose-dead.’ Japanese has a term colexifying ‘nose’ and ‘snot’ directly, as does Bislama, which also has the optional redundant term *doti blong nus* ‘rubbish/pus poss nose.’ Furthermore, there are semianalyzable terms with the identifiable meaning ‘nose’ in Highland Chontal, San Mateo del Mar Huave, Kashaya, Kiliwa, Pipil, and Rotuman. Finally, Hausa, Khoekhoe, Koyraboro Senni, Sora, Wintu, and Kapingamarangi have terms that are either ambiguous between nominal ‘snot’ and verbal ‘to blow the nose’ (Hausa, Koyraboro Senni, Wintu, Kapingamarangi), or derived from verbs with that meaning (Khoekhoe, Sora). Similarly,

the Acoma term is derived from a verb meaning ‘to have a runny nose,’ and ‘snot’ and ‘runny nose’ are colexified directly in Waris, Cheyenne, Rotuman, and Tetun.

Other associations include: Efik *mk’pö* is also a variant form of *ñk’pö* ‘thing, utensil, event, cause,’ Hausa *majina* also means ‘face,’ Buin *kina* colexifies ‘coconut sprout’ and ‘kina’ (the currency of Papua New Guinea), Burarra *an-gulol* is analyzable as ‘CLASS.MASCULINE-semen/rotten,’ and Sko *lóeri* might contain *loe* ‘ear.’ Bezhta *xida* also means ‘dew,’ and Japanese *hana*, with different prosodic structure, also means ‘flower.’ Carrier *nêninthastles* contains *ni* ‘nostril’ and *hwotles* ‘mud.’ Chickasaw colexifies ‘snot’ with “growth on a turkey’s beak,” while Pawnee *piruus* also means ‘be crooked, bent.’ Aguaruna *búshuk(u)* is also the name of a species of edible mushroom, while Cubeo *cōenó* also denotes the concept ‘tar,’ Ancash Quechua *puqru* also means ‘abscess,’ and Wayampi *ami* also ‘deceased.’ Hani *aqbeil beilgaol* ‘liquid nasal discharge’ contains *gaol*, meaning ‘clean water’ inter alia, and *aqbeil beilniul* ‘dense or solid nasal discharge’ contains *niul*, meaning ‘green’ inter alia. Bwe Karen *nekhà’i* might contain *khà’i* “a variety of edible wild fern” plus a prefix for body parts. Mandarin *ti4* also means ‘to cry,’ ‘tear,’ or ‘dripping’ (reflecting the same Early Middle Chinese etymon, Pulleyblank 1991: 305).

### 137. *The Semen*

Representation: 45%

Motivated: 46.7%

Thereof Analyzable: 26.8%

Thereof Colexifying: 21.5%

Thereof by Contiguity: 9.8%

Thereof by Similarity: 18.4%

Recurrent associated meanings: water/juice, seed, penis, pus, child, candle, egg, testicle, white, brain, resin, milk, urine

Quite frequently, terms for ‘semen’ are complex of the lexical type, with one constituent meaning ‘penis.’ As for the second element, ‘water,’ ‘liquid,’ and/or ‘juice’ is most common, as in Kosarek Yale *kede mak* ‘penis water/juice.’ Such terms are also found in Kyaka, Muna, Cayapa, Yanomámi (where ‘penis’ is colexified with ‘seed’), and Tetun. Meyah and Hawaiian directly colexify ‘water, juice’ (Hawaiian also “liquids discharged from the body” generally and Meyah also ‘river,’ compare section 47 for this pattern) with ‘semen,’ Mbum has *mbii gûn* ‘water child,’ Rotokas *ruve ovi* ‘slimy CLASS.LIQUID,’ which also means ‘slime,’ Guaraní *kuimba’e rykue* /kuimba’e tykue/ ‘man juice,’ and Bislama has *waet-wota* ‘white-water’ (for the association with ‘white,’ compare also Kiliwa *?-l-sap* ‘DN-ILL+white’ and Hawaiian *keakea*, with the reduplication base *kea* ‘white’). Nez Perce has *simqé-heqs* /símqe-heqs/ ‘penis-pus,’ and ‘semen’ is colexified with ‘pus’ in Ket, Lake Miwok (which also has a complex redundant term), Carib, and Yay. In Ket and Lake Miwok, ‘brain’ is furthermore colexified. Toaripi has *fe-oro* ‘penis-lime,’ and semianalyzable terms where the identifiable constituent is ‘penis’ are found in Kaluli, One, and Yir Yoront.

Khoekhoe colexifies ‘semen’ with ‘urine’ by an archaic term. Similarly, Tuscarora has *u?nhęhsú?kri?*, analyzable as /u-?nhęhs-ukr-i?/ ‘NOUN.PREFIX-egg/testicle-rubbish-NOUN.SUFFIX,’ with the stem *-?nhęhsukr-* also yielding *u?nhęhsú?kreh* “foul or disgusting urine, a slovenly or slatternly person so filthy as to emit an odor of urine.” Associations

with ‘egg’ are also found in other languages: Rama has *yáat aríra* ‘testicle/egg string/fold’ (‘testicle’ and ‘semen’ are colexified in Guaraní), and Kiowa directly colexifies ‘egg’ with ‘semen,’ as well as with ‘child’ (there is also a redundant complex term on the basis of this term with the additional constituent meaning ‘white’). This latter association with ‘child’ is also found in Wintu, Wayampi, and Hawaiian, and, as already noted above, Mbum has *mbii gân* ‘water child;’ compare also Guaraní *ta’yi* /*ta’y-i*/ ‘son/clot-DIM.’ Miskito and Hawaiian colexify ‘semen’ with ‘sap, resin.’ In fact *wai*, the relevant Hawaiian term has very broad reference, including ‘liquid’ generally, but also ‘honey’ and any liquid discharged from the body inter alia.

Dongolese Nubian, Basque, Greek, Welsh, Wintu, Aymara, and Guaraní colexify ‘semen’ with ‘seed’ (the Basque term also means ‘breed’ and has a verbal reading ‘to raise, grow’ inter alia, and the Dongolese Nubian term contains an element meaning ‘to sow’), and San Lucas Quiaviní Zapotec and Bislama with ‘milk.’ Cubeo (by a term whose root suffixed with different classifiers yields ‘gasoline lamp,’ ‘lantern,’ and ‘battery’), Macaguán (by the term *pitiócha*, analyzable as /*pe-itiót-ja*/ ‘3SG.POSS-enlighten-NMLZ’), and Sáliba colexify ‘semen’ with ‘candle.’

Other associations include: Hausa has *zuwan kai* for both ‘semen’ and ‘orgasm.’ *Zuwa* is glossed as ‘coming,’ and *kai* has extremely many meanings, among them ‘to come, arrive,’ and ‘head, top.’ Khoekhoe colexifies ‘semen’ with ‘egg of frog’ or ‘egg of fish.’ The Yoruba term *àtò* contains *tò* ‘to discharge liquid,’ while Burarra *burpur* also means ‘mould,’ and *gu-lol* gives rise to *-gulol* ‘rotten.’ Rotokas colexifies ‘semen’ with ‘spit.’ Ket *do’ŋ* is also the numeral ‘three,’ in Blackfoot, ‘semen’ is *okoyiim*, literally ‘his wolf,’ Lake Miwok *póta* also means ‘foam’ and ‘be gray, be cloudy’ (there is the optional complex term *?eláyni póta*, literally ‘children foam,’ for ‘semen’), Nez Perce *cép* also means ‘arrow,’ ‘bullet,’ and ‘cannon ball.’ Wintu *kur* also means ‘to be born,’ ‘to bear a child,’ ‘to be fertile,’ and San Lucas Quiaviní Zapotec *mo’c* (< Span. *moco*) also ‘phlegm.’ Arabela *mashiquia* also denotes ‘brushwood flowing downstream when the river is rising,’ and Carib *apy tano* contains *apy* ‘loin.’ Guaraní *kuimba’e rykue* contains *kuimba’e* ‘man,’ and Lenakel *nipiknisi* appears to be analyzable as /*nipik-nisii*/ ‘tail-excrement-’ (perhaps *nipik-* is also a colloquial designation for ‘penis,’ compare evidence in section 62 from other languages).

### 138. *The Stomach*

Representation: 97%

Motivated: 32.5%

Thereof Analyzable: 9.2%

Thereof Colexifying: 23.6%

Thereof by Contiguity: 5.8%

Thereof by Similarity: 4.7%

Recurrent associated meanings: guts/innards, womb, heart, eat/food, pregnancy, inside, crop, liver, feel/think, front, bag, middle, bladder, waist, faeces, chest, hypochondria, navel, diarrhea, liquid/water, big

Khoekhoe, Mbum, Kwoma, Ngaanyatjarra, Southeastern and Western Tasmanian, Badaga, Basque, Khalkha, Laz, Cheyenne, Chickasaw, San Mateo del Mar Huave, Kiliwa, Lake Miwok, Nez Perce, Wappo, Guaraní, Miskito, and Hawaiian colexify ‘stomach’ (or ‘belly,

abdomen') with 'guts' and/or 'innards' more generally (note also Pipil *-tu:xih* 'stomach' in the Cuisnahuat dialect and *tuxih* 'intestines' in the Santo Domingo de Guzmán dialect, that Burarra *-buka* also denotes "part of intestines," as well as the presence of this association in Indo-European, mostly Slavic, Buck 1949: 254). The association is realized formally by noun class alternation in Swahili and by gender alternation in Embera. Ngambay has *kéy bò sìn* 'house big guts' (note also the redundant Kyaka term *anda-romba* 'house-belly/stomach'), and Bororo has *peguru kodobo-reu* 'guts type.of.basket-like,' alongside a term directly colexifying 'faeces,' as well as 'rest, residue' with 'stomach.' An association with 'faeces' is also found in Tsafiki (*hua pe-coló* 'big excrement-package'), and a semianalyzable term with 'faeces' is also found in the related language Cayapa. Piro has *hitška-mapa* 'excrement-bag/bladder' (compare the colexification of 'stomach' and 'bladder' in Nuuchahnulth and Hawaiian), while Hawaiian and Rotuman colexify 'stomach' with 'bag' among other meanings; this association is also found in some Indo-European languages, in particular Celtic and Germanic (Buck 1949: 253). Katcha, Yoruba, Kwoma, Toaripi, Badaga, Khalkha, Welsh, Carrier, Lesser Antillean Creole French, Nez Perce, Yaqui, Bislama, Hawaiian, and Rotuman colexify 'belly' and/or 'stomach' with 'womb' (also very common throughout Indo-European, Buck 1949: 252), and similarly, Buli, Rendille, Ngaanyatjarra, and Bislama with 'pregnancy' and/or 'be pregnant.' 'Stomach' and 'heart' are colexified in Ngambay, Abzakh Adyghe, Badaga (where the gloss for 'heart' is in quotation marks, suggesting that this usage may be figurative), Laz, San Mateo del Mar Huave (by the analyzable term *omeaats-aran* 'inside-INAL.POSS'), Itzaj (by the analyzable term *pusik'al* /*puus-ik'-al*/ 'dusting-wind-COLL'), Santiago Mexquititlan Otomí, Yuki, Arabela, and Guaraní. Analogously, Kiowa has *t'ejin-t'qu* 'heart-water' ('liquid' and 'stomach' are colexified in Bororo among other meanings, and there is a semianalyzable term featuring a constituent meaning 'water' in Kashaya). A complex term for 'stomach' on the basis of 'heart' is also found in Breton (Buck 1949: 253). Ngambay, Khalkha, and Yanomámi colexify 'stomach' with 'liver,' and Malagasy has *ambavafô*, containing *vava* 'mouth' and *fô* 'heart;' note that there is a diachronic connection between 'stomach' and 'mouth' in Greek. Buin and Lesser Antillean Creole French colexify 'chest' (Buin 'chest of man' more specifically), and Basque 'bosom, breast' alongside 'gulf, bay.' Basque also colexifies 'belly, paunch' with 'navel,' and in Kyaka, there is a term that varies dialectally between these meanings. In Khoekhoe, there is a general term referring to internal organs of the trunk, including the 'stomach.'

Abzakh Adyghe and Itzaj colexifies 'stomach' with 'middle' (Baruya less generally with 'middle of the body,' Arabela similarly with 'center,' and Cashinahua with 'thick part in the center of something'), and Cavineña has *e-care-nani* 'INAL.POSS-middle/waist-hole.' The abstract notion 'inside' is colexified in Hausa, Dongolese Nubian, and Lake Miwok (note also the derived term in San Mateo del Mar Huave just mentioned above). Mirroring the Cavineña association with 'waist,' Cahuilla and Samoan 'stomach' with 'waist' directly.

Wintu colexifies 'stomach' with 'wrinkles' (and perhaps 'honeycomb tripe'), and there is a semianalyzable term where the identifiable constituent means 'wrinkled' in Wayampi. Lake Miwok, Arabela (by a semianalyzable term containing the classifier *-co* for receptacles), Wayampi, Yanomámi, and Hawaiian colexify 'crop,' Ngaanyatjarra, Yir

Yoront, and Samoan ‘front (of person),’ and Basque and Bororo ‘hypochondria.’ Associations with ‘food’ or ‘eating’ are found in Oneida (*-khwálákhwa?*, containing the roots *-khw-* ‘food’ and *-l-* ‘be in or on’), Kaingang (*jên né* ‘eat container’), Tehuelche (*xa:t'en ~ xat'en*, derived from *xa:t'e* ‘to eat’ and colexifying ‘food.’ On the basis of this term, there is the redundant term *xat'en p'ate* ‘stomach/food skin;’ a semianalyzable term with ‘skin’ as the identifiable constituent is found in Berik), and Hawaiian (*pu'u 'ai-'ai* ‘protuberance eat/food-RED’). Furthermore, Yir Yoront has the respect vocabulary term *maymay* reduplicated from *may* ‘food.’ Finally, Aguaruna and Rotokas colexify ‘stomach’ with ‘diarrhea’ (Aguaruna also with ‘cholera’), and in Kwoma, Ngaanyatjarra and Badaga, the ‘stomach’ is also considered the seat of emotions.

Other associations include: Buli *puuk* also denotes a particular clay vessel, and, in addition, colexifies ‘flower, blossom’ and ‘foam, lather.’ Hausa *ciki* also denotes the strips of cloth in a garment, and “a children’s aquatic game.” Khoekhoe features a term with very broad reference to ‘innards’ and ‘offals,’ including alongside ‘heart’ and ‘liver’ also ‘lungs’ and ‘kidneys;’ it also means ‘interior’ generally. Rendille *úur* also denotes the ‘character’ of a person. Baruya *munya* also means ‘above, high, top’ as well as ‘stone adze.’ Buin *moo-nogu* is analyzable as ‘dirt-like.’ The Dadibi term *moni hamago* contains *moni* ‘large,’ and Kyaka *romba* also denotes the “outer skin and fat layers over stomach.” Muna *taghi* is also used to refer to ‘lees’ or ‘sediment,’ and *randa* can also mean ‘bruised’ (also of fruits). Ngaanyatjarra *tjuni* also is used to refer to the “hollow of anything concave” and the “round part of a fruit,” while Nunggubuyu *mulgu* also denotes the ‘stomach lining.’ Rotokas *kovapato* appears to be derived from *kova* ‘growth,’ ‘to grow, mature.’ Toaripi *luka* is also the name of a tree with hard timber, and Yir Yoront *pirm* may also mean ‘body cavity’ as well as ‘the interior of house.’ Abzakh Adyghe *gʷə* also means ‘surface, territory,’ Badaga *bevaru* is also a verb meaning “to spread grain out, to sort out grains,” and *karu* also means ‘dysentery’ and denotes “something that came from the stomach” inter alia. Basque *urdail* also means ‘rennet,’ Ket *huŋ* also ‘room,’ and Khalkha *gedesy(n)* is also the name for “the loop for o/u in Khalkha script” and may be derived from *gede* ‘back of neck, occiput.’ Haida *k'iji* also denotes the ‘swim bladder of fish.’ Welsh colexifies ‘calf,’ and Lesser Antillean Creole French *vant* also means ‘sale, auction.’ Tuscarora *útkweh* also means ‘groin’ and ‘pleura,’ and Central Yup'ik colexifies ‘gizzard.’ Arabela *sara-ca* appears to be analyzable as ‘uvula-CLASS.FRUIT.’ Chayahuita *anpopi-tě* is analyzable as ‘pith-CLASS.INSTRUMENT.’ The relevant Cubeo term shares its root with terms for ‘sweet potato’ and ‘alcoholic beverage made from sweet potato,’ and Guaraní colexifies ‘stomach’ with ‘spirit, conscience.’ Lenakel *tipweua* is also the name of a kind of breadfruit. Rotuman *taga* can also refer to a ‘pocket’ inter alia, while Samoan *manava* also means ‘smooth, soft side of a thing’ and, figuratively, ‘child.’ Sedang *potok* is also used to refer to the ‘abdomen’ of insects.



139. *The Sweat*

Representation: 75%

Motivated: 17.7%

Thereof Analyzable: 7.7%

Thereof Colexifying: 10.1%

Thereof by Contiguity: 11.4%

Thereof by Similarity: 3.2%

Recurrent associated meanings: heat/warmth, water/liquid/juice, steam

The most common of the relatively few lexico-semantic associations for 'sweat' (or 'perspiration') is that with 'heat/hot,' 'heated,' or 'warmth/warm' (as it is in Indo-European, Buck 1949: 263-264). These meanings are colexified in Buli, Hausa (alongside other apparently unrelated meanings), Badaga, Ket, Abipón, Aguaruna, Huambisa, Miskito, Fijian, and Samoan. Semianalyzable terms where a constituent with this meaning is present are found in Kolyma Yukaghir and Ineseño Chumash, Oneida has the verb *-ate?tukhwálha?* 'to sweat' revolving around the root *-?tukhwal-* 'to be hot' and Kiowa has *shdl-t'əy* 'be.hot-water.' Unsurprisingly, Kiowa is not the only language to feature terms for 'sweat' related in some way to 'water,' 'liquid,' or even 'juice.' In fact, Toaripi has *maea ma ma* 'body water RED,' Abzakh Adyghe *pš'ent'e ~ pš'ant'e* contains elements meaning 'water' and 'violent,' Carrier has *nê-tsi-n-thû* '??-head-EPEN-water,' Santiago Mexquititlan Otomí *nxanthe /xani-dehe/* 'sprinkle-water,' which colexifies, like White Hmong, 'steam,' and Pawnee *kickaasisu* is analyzable as */kic-ka'as-his-u'/* 'liquid-break.out-PERF-NOM,' Lesser Antillean Creole French directly colexifies 'sweat' with 'water' (as well as 'rain'), and semianalyzable terms containing an element 'water' and/or 'liquid, juice' are found in Piro and Tsafiki.

Other associations include: Hausa *ji'bi* can also refer to a "large quantity," Kaluli *ha:fof* might be related to *ha:fó*: "deep breathing, panting, wheezing, asthma," Kyaka *pusi* also means 'cat' due to borrowing from Tok Pisin, and Ngaanyatjarra *parlulungu* means 'humidity, humid' and in the Northern dialect also 'sweat.' Kosarek Yale *wihin* 'sweat, tiredness' contains *wihi* 'ripe, big; tired.' Yir Yoront has *morr-ninn* 'body-sweat.smell.' Badaga *uri ~ huri* inter alia also means 'to flare up,' 'to fry,' 'jealousy,' and 'venom.' Basque *izerdi* also means 'dampness' and 'work,' while among the meanings colexified by Khalkha *kəlysy(n) ~ kəlesy(n)* is also 'payment for work, hire, fees.' Wintu *kiw* also denotes 'cooking stones' and a river rock which is put into acorn soup when cooking. San Mateo del Mar Huave colexifies 'sweat' with 'dermatitis,' while Central Yup'ik *uquryak*, a Hooper Bay and Chevak dialect term for 'heavy sweat' is analyzable as */uquq-yak/* 'oil-thing.similar.to.' Cashinahua *dabixtun* is also used to refer to 'grease on the body of a newborn.' Guaraní *ty'ái* appears to be analyzable as */ty-ai/* 'urine-line.' Sáliba *aixito* denotes also 'hot, boiling water' specifically, while Wayampi *pili?ay* is analyzable as */pili-ay/* 'smell-bad.' Bislama *swet* (< Engl. *sweat*) also means 'to put a lot of effort into something,' Kapingamarangi *hee* also 'where,' and Hawaiian *hou* also means 'new, fresh' and 'to push, thrust' inter alia. Samoan colexifies 'sweat' with 'wither' and 'waterfall.' Finally, Yay *haan*<sup>5</sup> also means 'promise.'

140. *The Tear*

Representation: 86%

Motivated: 41.6%

Thereof Analyzable: 37.8%

Thereof Polysemous: 3.8%

Thereof by Contiguity: 34.2%

Thereof by Similarity: 1.2%

Recurrent associated meanings: water/liquid/juice, eye, cry

Very frequently, terms for ‘tear’ are characterized by analyzability of the lexical type, with constituents meaning ‘eye’ (which sometimes colexifies further meanings such as ‘face’) and ‘water,’ ‘liquid’ and/or ‘juice,’ as in Itzaj *k’a’ ich* ‘juice/liquid eye.’ Such terms are featured in Buli, Mbum, Berik, Abzakh Adyghe, Badaga, Ket, Sora, Kolyma Yukaghir, Carrier (where additional constituents are present: *nê-na-tsel-thû* is analyzable as ‘human-eyes-anus-water’), Upper Chehalis, Chickasaw, Haida, San Mateo del Mar Huave, Kashaya, Kiliwa, Lesser Antillean Creole French, Santiago Mexquititlan Otomí, Pawnee (in this language, the term is identical on the surface with that for ‘jelly, jam,’ but has a different underlying morphology), Pipil, Wintu, Yuki, Copainalá Zoque, Arabela (where the term is more precisely not of the lexical type, but derived by a classifier for liquids), Bororo, Embera, Guaraní, Hupda, Jarawara, Maxakalí, Miskito, Rama, Tsafiki, Wayampi, Bislama, Fijian, Hawaiian, Lenakel, Malagasy, White Hmong, Sedang, Tetun, and Yay. The pattern is furthermore etymologically detectable in Yoruba. There are, however, also other complex terms involving a constituent meaning ‘eye.’ Mbum has *ôî-yâr* ‘at-eye,’ Toaripi *ovo-roro* ‘eye-rubbish,’ Yuki *hul-kat* ‘eye-wet,’ Rotuman *sui ne mafa* ‘bone GEN eye’ (the *sui* in this expression is thought to be a mere “doublet” of *sui* ‘bone’ by the lexicographer), and Bororo has an alternative term which, alongside constituents meaning ‘eye’ and ‘water,’ also features elements meaning ‘thing’ and ‘fire.’ Semianalyzable terms where the identifiable constituent is ‘eye’ are moreover present in Kanuri, Sentani, Yir Yoront, Highland Chontal, Kiowa, Xicotepec de Juárez Totonac, Tuscarora, Wintu, Kaingang, Sáliba, Hani, and Samoan.

Likewise, there are complex terms in some languages in which one constituent is ‘water,’ ‘liquid,’ or ‘juice,’ but the other one is not ‘eye.’ In this case, a verb meaning ‘to cry’ is the most frequent alternative. San Lucas Quiaviní Zapotec, for one, has *nnyí’ihs rùàa’n* ‘water cries,’ and such terms are also found in Efik and Yuki, while Bora has *máátyo-u* ‘crying-CL.round’ and Chayahuita *na’nëi’*, analyzable as /na’nërin-i’/ ‘cry-CLASS.LIQUID.’ Alternatively, Dongolese Nubian has *ôñmisse* /ôñ-mísse/ ‘cry-sprinkle,’ Koyraboro Senni, Ngambay, Buin, Upper Chehalis, and Aymara colexify the relevant meanings, and Kwoma has a semianalyzable term where the identifiable constituent is ‘cry.’

Other complex terms involving ‘water,’ ‘liquid,’ or ‘juice’ include Khoekhoe has *llgam-ro-s* ‘water-DIM-3SG.FEM’ and Wappo *hu-méy* ‘head-water.’ Ancash Quechua and Manange colexify the relevant meanings, and semianalyzable terms where the identifiable constituent means ‘water,’ ‘liquid,’ or ‘juice’ are present in Kosarek Yale, Cayapa, Kaingang, Yanomámi, and Bwe Karen.

Other associations include: Hausa *k’walla* also means “repletion with fura” (a kind of dish), Koyraboro Senni *heeni* also means ‘to cry’ and “to make any loud or continuous

noise” in general (an association also attested in Indo-European, Buck 1949: 1130-1131), Rendille colexifies ‘tear’ with ‘seeds,’ Kyaka with ‘garden,’ and Basque *negar* also means “lamentable, deplorable” and “dripping.” *Malko* is also the name for a kind of small pear. Bezhta *maq’o* also means ‘loom.’ The Khalkha term *nilbusu(n) ~ nilmusu(n)* also may refer to ‘mucus’ or ‘spittle,’ Abipón *-aci-* also means ‘ashes’ and a root of the same shape also occurs in the term for ‘tongue.’ Carib *-enakulu* also means ‘dirt in eye.’ Cavineña *paanacaca* contains *caca* ‘little,’ Toba colexifies ‘tears’ with ‘lung,’ and Lesser Antillean Creole French *lam* also means ‘spirit, soul,’ presumably because of phonological collapse of Fr. *larne* and *l’âme*. Finally, Hawaiian *pūkai* ‘lime bleach for hair, to bleach’ rarely assumes the meaning ‘salty tears;’ this term may be related to *kai* ‘sea.’

#### 141. *The Tendon*

Representation: 75%

Motivated: 70.2%

Thereof Analyzable: 5.7%

Thereof Colexifying: 64.5%

Thereof by Contiguity: 4.5%

Thereof by Similarity: 0%

Recurrent associated meanings: vein/artery, nerve, muscle, thread/twine/fibre, root, gristle, string/cord, bowstring, line, flexible thing/elastic, sword, flesh

Very frequently, languages colexify ‘tendon’ with ‘vein’ and/or ‘artery’ cross-linguistically judging from the evidence of the sample. This is the case in Bakueri, Buli, Efik, Hausa, Dongolese Nubian, Yoruba, Baruya, Berik, Buin, Dadibi, Gurindji, Kyaka, Muna, Ngaanyatjarra, Nunggubuyu, One, Toaripi, Sahu, Kosarek Yale, Yir Yoront, Chukchi, Khalkha, Nivkh, Sora, Itzaj, Highland Chontal, Haida (for younger speakers only), Lakshota, Itzaj, Xicotepec de Juárez Totonac, Tuscarora, Arabela, Cashinahua, Cavineña, Cayapa (by a semianalyzable term containing an element meaning ‘line;’ ‘line’ and ‘tendon’ are colexified in Japanese and Biloxi), Chayahuita, Guaraní, Lengua, Miskito, Tehuelche, Toba, Yanomámi (by the analyzable term *mathōyāhi* /*matha-yāhi*/ ‘leg-flesh,’ compare Tetun *na’an-isin* ‘meat/flesh-flesh’ and note also that Kyaka and Toaripi have redundant complex terms with ‘flesh’ as the meaning of the additional constituent to single out the meaning ‘tendon,’ and that there is a semianalyzable term with the identifiable constituent meaning ‘muscle, meat’ in Hani), Bislama, Fijian, Great Andamanese, Hawaiian, Lenakel, Malagasy, White Hmong, Rotuman, Takia, and Yay. Moreover, Bora has a complex term for ‘tendon’ on the basis of ‘vein’ featuring, among other additional constituents, *méjpi* ‘body,’ Embera has *hir’ūkēḡgúdroma*, containing *hir’ū* ‘foot’ and *kēḡgú* ‘nerve, vein,’ and Samoan has *uaua i so’oga* ‘vein/artery/pulse LOC joint.’ Furthermore, there is the term *otsinuhyáhta?* *ahsli:ye* in Oneida, containing *otsinuhyáhtu?* ‘vein’ and *ahsli:ye* ‘string, thread, yarn;’ however, the term for vein is itself semianalyzable, containing the root *-nuhy-* for ‘sinew.’

Many of the languages just mentioned, but also some others, colexify ‘tendon’ also with ‘nerve.’ This is the case in Buli, Hausa, Toaripi, Yir Yoront, Chukchi, Khalkha, Sora, Carrier, Ineseño Chumash, Itzaj, Lakshota, Nez Perce, Xicotepec de Juárez Totonac, Arabela, Aymara, Bororo, Cavineña, Chayahuita, Guaraní, Miskito, Ancash Quechua, Tehuelche, Yanomámi (again by the analyzable term *mathōyāhi* /*matha-yāhi*/ ‘leg-flesh’), Hani, Ha-

waiian, and Rotuman (colexifying ‘large nerve’ specifically). Furthermore, Muna, Ngaanyatjarra, Biloxi, Upper Chehalis, Ineseño Chumash, Wappo, Guaraní, Jarawara, Lengua, Wichí, Yanomámi, Fijian, Hawaiian, Lenakel, Malagasy, and Tetun colexify ‘tendon’ with ‘muscle’ (Yanomámi and Tetun more specifically with ‘muscular tissue’), Cheyenne, Haida, Itzaj, Lake Miwok, and Wappo with ‘gristle,’ and Aymara and Rotuman with ‘flexible thing’ and/or ‘elastic’ respectively (similarly, Ancash Quechua *anku* also means ‘hard, difficult to break by stretching’).

There is a further association, namely colexification with ‘root,’ which is found in Kwoma, One, Chickasaw, Ineseño Chumash, Pawnee, Jarawara, Lengua, Miskito, and Hawaiian (colexifying ‘small root, rootlet’ specifically). Gurindji colexifies ‘single root of tree’ specifically; furthermore, Kiliwa has a derived term, and there is a semianalyzable term where the identifiable constituent means ‘root’ in Highland Chontal. Khoekhoe, Tasmanian (Northeastern, Middle-Eastern, and Southeastern), Chukchi, Khalkha, Kolyma Yukaghir, Biloxi, Wintu, Yuki, Central Yup’ik, Lengua, and Bislama, by functional or provenience contiguity, colexify ‘tendon’ with ‘thread,’ ‘twine,’ and/or ‘fibre,’ Bislama also with ‘fishing line’ specifically. Finally, Upper Chehalis and Ineseño Chumash colexify ‘bowstring,’ Dongolese Nubian, Biloxi, Wintu, Copainalá Zoque, and Bislama ‘string’ more generally or ‘cord,’ and Japanese and Mandarin Chinese ‘sword’ (the association was borrowed along with the relevant term *jian*<sup>4</sup> from Chinese into Japanese).

Other associations include: Hausa *jijiya* also denotes a “tightening string of a drum,” and Baruya *wirila* is also used to refer to “ridges on a shield” and “stringy fibres of wild yam.” Muna *ue* is also denotes a unit of measurement inter alia. Ngaanyatjarra *pulyku*, in the Northern dialect, also denotes the “Pencil Yam.” Abzakh Adyghe *le*, only perhaps meaning ‘tendon,’ otherwise also means ‘be capable’ and ‘white, light.’ Basque *zurda* means ‘mane’ and ‘fishing-line’ inter alia, and in the dialect of Lapurdi also ‘tendon.’ Khalkha *xujang* also means ‘rheumatism, arthritis.’ Haida *xay* also means ‘warp,’ Itzaj *jich’* also ‘tighten’ (another Itzaj term, *xich’el*, consists of a marker for inalienable possession as well as the root *xich’*, which can also mean ‘wiry, thin’), and Nuuchanulth *tuktapt* is also the name of a spirit. Xicotepec de Juárez Totonac has a semianalyzable term containing an element meaning ‘bone.’ Tuscarora *un̥ghyáθeh* also means ‘plantain,’ Wintu *la’* also ‘skin’ and ‘stretch,’ and Copainalá Zoque *ληληληλ* also ‘rope.’ Aguaruna *wánus(e)* also means ‘ankle bone,’ Guaraní *tajygue* also ‘strength’ and ‘lungs,’ and Jarawara *kowisa-ri* appears to be analyzable as ‘hurt-NMLZ,’ while *habi* also colexifies ‘tuber,’ and, figuratively, ‘courage’ and ‘audacity.’ Imbabura Quechua colexifies ‘tendon’ with ‘vine.’ Bislama *string* also means ‘erection, to have an erection,’ and Fijian *ua* also ‘wave, tide.’ Figuratively, Hawaiian *a’a* also means ‘womb, offspring’ (similarly, Lenakel *nouanul-*, containing *noua-* ‘mouth, opening,’ ‘fruit’ colexifies ‘genealogical line, family, descent-group’) as well as “to send greetings of love, joyous hospitality.” Hani *saqquq* also denotes the ‘pulse of a vein.’ Finally, Sedang has *tróang húan* ‘road grow/sprout/bud,’ and Yay colexifies ‘tendon’ and ‘vein’ with “violin, fiddle.”

142. *The Testicle*

Representation: 84%

Motivated: 47.4%

Thereof Analyzable: 21.4%

Thereof Colexifying: 26.0%

Thereof by Contiguity: 10.3%

Thereof by Similarity: 11.9%

Recurrent associated meanings: egg, scrotum, seed, round, fruit, ball, child/offspring, stone, penis, genitals, kidney, nut, energy

The 'testicles,' when expressed by motivated terms, are most frequently associated lexically with small roundish objects by metaphorical transfer. The findings on the basis of the present sample are largely in line with the results arrived at by Brown and Witkowski (1981) and Buck's (1949: 257) brief statement on Indo-European in particular, though some associations not present in Brown and Witkowski's study are uncovered here.

A frequent transfer pattern is that from 'seed' to 'testicle.' Fijian has *sore-ni-qala* 'seed-POSS-scrotum,' and derived terms, such as Chickasaw *intalop* /im-talop/ 'DAT-seed,' are also featured in Ineseño Chumash, Arabela, and Hawaiian. One has *amplu tala* 'male.genitals seed/round.thing,' while Welsh, San Mateo del Mar Huave, San Lucas Quiaviní Zapotec, Guaraní, Toba, Hawaiian, and Kapingamarangi colexify 'seed' and 'testicle' directly (Kapingamarangi 'seed of breadfruit' and Welsh 'pit of fruit' more precisely; Toba also colexifies 'bullet'), and there is a semianalyzable term in Kaingang. Like Kapingamarangi, an association with 'nut' and 'walnut' more particularly is present in Khalkha.

Moreover, White Hmong has *noob-qes* 'seed-egg,' and indeed, the association between 'testicle' and 'egg' is also common cross-linguistically. Sora, for one, has *'arre:-'kad-ən* 'egg-male.genitals-N.SFX.' An analyzable terms of the lexical type (with 'scrotum' acting as the contiguity anchor) is found in Dongolese Nubian, and derived terms are featured in Basque (where the 'testicle'-word is the plural of that for 'egg'), Blackfoot and Lengua, semianalyzable terms perhaps in Hupda, Macaguán, Lenakel, and Manange, and direct colexification occurs in Efik, Nganyatjarra, Carrier, Highland Chontal, Itzaj, Santiago Mexquititlan Otomí, Pawnee (with slight phonological deviations), the Cuicahuat dialect of Pipil, Tuscarora, Bororo, Miskito, Piro, Ancash and Imbabura Quechua, Rama, Tsafiki, Yanomámi, Bwe Karen, and Lenakel; the association is diachronically recoverable in Wayampi. Moreover, Berik has *bol sui* /bola sui/ 'ball egg,' and similar associations on the basis of 'ball' are found in Japanese (*kin-tama* 'gold-ball'), Mandarin (*gao1-wan2* 'swamp/high/eminent/praise-globe/ ball/lump'), and Vietnamese (*hòn dái* 'ball genitals'), while 'ball' and 'testicles' are colexified in Nuuchahnulth and Bislama (compare the situation in English). Lenakel features a semianalyzable term where the identifiable constituent *noua* means 'fruit,' and associations with 'fruit' specifically are also found in Toaripi (*kō fare* 'scrotum fruit,' with *kō* also meaning 'breast, milk'), Samoan (*fuā-manava* 'fruit-belly'), and Tetun (*lasan-fuan* 'penis-fruit/heart'). 'Fruit' and 'testicle' are colexified in Yir Yoront (also with 'rolled string'), Cavineña, Bislama, and Hawaiian. A term bearing an association with 'stone' is featured in Ineseño Chumash (*is-xip* 'one's.own-stone'), a similar term is featured in Oneida, the association is realized by colexification in Welsh, and a

semianalyzable term where the identifiable constituent bears the meaning ‘stone’ is featured in Santiago Mexquititlan Otomí (note also the redundant Kanuri term *súsú tǎrwǎlǎm-bè* ‘stone male.testicles-of’). ‘Testicle’ and ‘kidney’ are colexified in Khalkha and Rotuman. In fact, alongside the associations with specific smallish round objects discussed so far, there are also languages in which terms for ‘testicles’ make reference to ‘roundness’ directly, as already foreshadowed by the One term mentioned above. Rotokas has *orikoroisi* apparently containing *roi* ‘intercourse’ and the classifier *isi* for round objects, Basque has *barrabil*, containing *bil* ‘round,’ Nuuchahnulth *hupkimt*, which is presumably analyzable as /hup-qimt/ ‘roundish.thing-CHUNK.SHAPED.OBJECT,’ Wintu *siw*, which is diachronically related to the word for ‘eyeball,’ Arabela, as alluded to above, has a term derived from ‘sap, juice’ by means of a classifier for round objects, Bislama *raon-wan* ‘round-one,’ Hani *davqwuv wuv-siq* ‘scrotum RED-CLASS.ROUND,’ and Vietnamese *tinh hoàn* ‘energy round.object’ (this association recurs in Bislama by colexification, where the relevant term is *paoa*, from English *power*). Hawaiian colexifies ‘testicles’ and ‘small round object’ directly, and Cayapa features a semianalyzable term with a constituent with precisely this meaning. As may have been noted from the discussion so far, ‘scrotum’ frequently acts as a contiguity anchor in complex terms of the lexical type. ‘Scrotum’ and ‘testicle’ are colexified furthermore in Buli, Koyraboro Senni, Swahili, Muna, Nunggubuyu, Aguaruna, Aymara, Cashinahua, Ancash Quechua, Bislama and Takia, and in Khoekhoe, the same root yields both ‘testicle’ and ‘scrotum,’ with different nominal designants disambiguating the referents. Another complex term involving a constituent meaning ‘penis,’ alongside the one in Tetun mentioned above, is Sko *ðebi*, perhaps /ðe-bí/ ‘penis-floor/shell,’ which colexifies ‘testicle’ with ‘cheek.’ Semianalyzable terms of this type are featured also in Kaluli, Sentani, Kosarek Yale, and Toba, while Tehuelche colexifies ‘penis’ and ‘testicle’ (for Sentani *mu haka*, where *mu* is ‘penis,’ compare *haka-bo-* ‘run away?’). Finally, figuratively, ‘testicles’ may also be extended to “son under three years of age” in Great Andamanese and to ‘offspring’ in Hawaiian, while Guaraní has *ta’yi* /ta’y-i/ ‘son/clot-DIM.’

Other associations include: Hausa *gwaiwa* also denotes “[t]he weight (clay or stone) at the end of the pole of an irrigation plant,” Ngambay *gèm* also ‘times,’ and Rendille *jiláh* also ‘burning coal.’ Anggor *timoeft* may also be capable of referring to a ‘tree root,’ and Kwoma *madii* also to “women’s protruding labia.” Badaga *ode* ‘animal testicle’ also means ‘to break’ and “wood, thicket, bush,” while Greek *órchis* also means ‘orchid.’ Kolyma Yukaghir *önd’ed-abut* is analyzable as ‘male-container,’ while Upper Chehalis *má’čari* also means ‘pear.’ Central Yup’ik *ingcu* is also used with the meanings ‘nosebead’ and ‘mantle in gas lamp.’ Bora *dómi-úúho* is analyzable as ‘pubis-CL.chunk,’ Tsafiki *pi’poca* might consist of *pi* ‘water, liquid, juice’ and *poca* ‘cane of guadúa bamboo,’ while Yanomámi *aruku* also denotes the ‘eggs and larvae of wasps and bees.’ Hawaiian *hua* also colexifies ‘tuber,’ ‘produce, yield,’ ‘ovum,’ as well as ‘word, figure,’ among other meanings, and the formally redundant reduplicate *huahua* also means ‘fruitful, productive, prolific, have many children, lay many eggs’ and is the name of a ‘vulgar gesture.’ Kowaū, another term in the same language, also means ‘fish eggs.’ Rotuman *ififi* can also be used with the meaning “in bunches or clusters.” Bislama *bol* (< Engl. *ball*) may also refer to ‘sac’ and the

“soft belly of coconut crab,” and *frut* also to “any individually sold item taken out of a packet.”

#### 143. *The Tongue*

Representation: 97%

Motivated: 20.6%

Thereof Analyzable: 3.1%

Thereof Colexifying: 17.5%

Thereof by Contiguity: 11.2%

Thereof by Similarity: 6.3%

Recurrent associated meanings: language/dialect/speech/word, blade, blade of oar, lick, trigger, foot of mollusk, mouth

‘Tongue’ is colexified with ‘language,’ ‘dialect,’ ‘speech,’ and/or ‘word’ in Hausa, Toaripi, Yir Yoront, Basque, Bezhta, Chukchi, Greek, Khalkha, Laz, Sora, Lesser Antillean Creole French, Quileute (here, the additional reading ‘language’ is rare), Cavineña, Embera, Rama, Bislama, and Hawaiian (this association is also common in Indo-European according to Buck 1949: 230, corroborating the commonness of the association in Eurasia detected in the sample, compare also Hilpert 2007), while Abzakh Adyghe has *bzeg* /bze-g°ə/ ‘language/area-surface,’ and Tasmanian (Northeastern) has a semianalyzable term where the identifiable constituent means both ‘mouth’ and ‘language’ (which are often colexified as well, compare section 124).

Otherwise, Wichí *tok’qjlhech’e* contains elements meaning ‘mouth’ and ‘egg,’ and semianalyzable terms where the identifiable constituent is ‘mouth’ are attested in Nunggubuyu, Northeastern Tasmanian, Kiliwa, Wappo, and Hupda (note also that Kashaya *ha’ba* is etymologizable as \*?aha-hiba ‘mouth-tail’). There is just one language to colexify ‘mouth’ with ‘tongue’ (as well as ‘manner of speaking’) directly, namely Miskito. Hausa, Basque, Fijian, and Malagasy colexify ‘tongue’ with ‘blade’ (similarly, Chukchi, Fijian, and Hawaiian with ‘blade of oar,’ and Nivkh and Arabela with ‘trigger’). Finally, Rama *kúup* ~ *múkup* contains *up* ‘eye,’ and Tsafiki *ni’caca* appears to be analyzable as /ni-ca’cá/ ‘seed-eye.’ Haida and Nuuchahnulth colexify ‘tongue’ with ‘foot of (a certain) mollusk,’ and Wintu and Manange colexify ‘tongue’ and ‘to lick’ (compare associations between ‘tongue’ and ‘lick’ in Indo-European, Buck 1949: 230).

Other associations include: Hausa *harshe* also means ‘flame’ and “[e]xtremity of a whip; point of a sword, knife, or loin-cloth.” Ngambay *ndòn* also means ‘to hunt,’ Anggor *tefū* also “inside of pandanus fruit,” and Burarra *ngarl* also denotes the ‘taste of fresh meat or seafood.’ Waris *minde* is also the name of a “pole with a hook for pulling fruit from a tree,” while Basque *mihi* also means ‘bit’ and ‘tap, flap’ inter alia. Basque *mingain* also means ‘masthead,’ Bezhta *mic* also ‘nettle,’ and Greek *glóssa* also denotes the ‘sole’ (*Solea solea*, compare German *Seezunge*). Ket *ēj* also means ‘river island’ and ‘pine,’ Khalkha *kele(n)* also ‘bell clapper,’ ‘tongue of buckle’ and “[a]nything resembling the tongue” generally, Japanese *shita* with different prosodic structure also ‘down, under,’ and Cheyenne *vétanove* also ‘tongue of wagon’ and ‘tongue of shoe.’ Haida *t’aangal* also means ‘barb of fishhook,’ and Itzaj *ak’* also ‘vine.’ Xicotepec de Juárez Totonac *sī’mā’kā’t* also means ‘uvula.’ Central Yup’ik *alungun* is the name of a certain dog-feeding trough, and is derived from *alunga*°

‘home-made dog food’ but also has the meaning ‘tongue’ in the Kuskokwim and Bristol Bay dialects. The root *-aci-* around which the Abipón term for ‘tongue’ revolves also means ‘tear’ and ‘ashes,’ in Cubeo, an identical root to that found in the term for ‘tongue’ yields the meaning ‘bundle’ when suffixed with another classifier, and Jarawara *abate* also means ‘cheek.’ Piro *n-u* may also be used with reference to a soft, fleshy object in general, while Wayampi *apekū* also means ‘reed of a clarinette,’ and Yanomámi *aka* is also used to refer to the ‘proboscis of the dipteros,’ as well as to the ‘starter of an outboard engine.’ Bwe Karen colexifies ‘tongue’ with ‘to poison,’ Lenakel with ‘fish,’ and the Hawaiian term *alelo ~ elelo* inter alia also denotes “meat of the sea egg or sea urchin,” *lelo* also “yellowish, especially the hue imparted to a whaletooth pendant ... by smoking,” and Samoan *laulaufaiva* contains *lau*, meaning either ‘leaf’ or ‘flat and thin object’ generally, as well as *faiva*, which can mean ‘fishing party,’ ‘job,’ ‘skill,’ and ‘business’ inter alia.

#### 144. *The Tooth*

Representation: 98%

Motivated: 22.5%

Thereof Analyzable: 3.1%

Colexifying: 19.4%

Thereof by Contiguity: 4.1%

Thereof by Similarity: 13.9%

Recurrent associated meanings: beak, cog, mouth, bite, claw of crab, tip/point, horn, thorn, ivory, jaw, sharp, blade, seed/grain of corn

Contiguity-based associations for this meaning (additional glosses such as ‘tusk’ or ‘fang’ are not reported in the below discussion) are relatively rare, and most recurrent patterns are metaphor-based extensions. Baruya, Gurindji, and Cahuilla colexify ‘tooth’ with ‘mouth,’ while Aymara has *laka ch'akha* ‘mouth bone.’ Moreover, Central Yup'ik has *keggun /kegge-(u)n/* ‘bite-device.for,’ such a derived term is also found in Kolya Yukaghir, while Abzakh Adyghe and Hawaiian colexify nominal ‘tooth’ with verbal ‘to bite’ (Abzakh Adyghe also more specifically with ‘to gnaw,’ this association is also etymologically recoverable in Indo-European, Buck 1949: 231). Kyaka and Jarawara colexify ‘tooth’ with ‘sharp’ (and Maxakalí *xox* is also glossed as “sharp fragment”). Ngambay and Wappo colexify ‘jaw’ with ‘molar tooth’ specifically, and Hausa and Dongolese Nubian colexify ‘tooth’ with ‘ivory.’

Turning to metaphorical extensions, Baruya, Kyaka, Itzaj, Arabela, Bora, Cashinahua, Lengua, and Rama colexify ‘tooth’ with ‘beak’ (Kyaka colexifies ‘narrow beak’ more specifically; note also Embera *kidhá* ‘tooth’ and *kidá* ‘beak,’ and compare section 5 for complex terms for ‘beak’ betraying this association), Basque, Greek, Khalkha, Welsh, Lesser Antillean Creole French, and Hawaiian with ‘cog,’ Kyaka, Chukchi, and Samoan with ‘horn’ (similarly, Takia *fai* colexifies ‘horn’ with “upper canine teeth” specifically, and also with ‘crocodile’), Toaripi, Bislama, and Hawaiian with ‘claw of crab,’ Kyaka, Kosarek Yale, and Lengua with ‘thorn,’ Lake Miwok and Jarawara with ‘blade,’ and Itzaj and Wintu with ‘seed’ or ‘grain of corn.’



Moreover, Toaripi, Kosarek Yale, and Jarawara extend 'tooth' to 'tip, point' in general (Toaripi also with 'summit' and 'corner'), and Baruya, Burarra, and Fijian colexify 'tooth' with 'edge.'

Other associations include: Efik *edet* also denotes a "gap in the edge of a tool" and "[t]he horns of a post between which a wall plate is laid." With the nominal designant -s, Khoekhoe (Haillom and Topnaar/ǀAonǀn dialects) *apu-b* assumes the meaning "empty ostrich shell" which is used as a vessel. Hausa *hak'ori* also means "rough surface," as well as "[a]n embroidered edging of garments." Ngambay *ngange* may also refer to a 'limit, border.' Anggor *yahafi* seems to contain *yah* 'to say,' and Buin *kompe* also denotes "the bony toothridge in the beak of a hornbill or other bird." Kwoma *pu* also means 'conch shell, conch shell trumpet.' Kyaka *nenge* also means 'food,' and Muna *wangka* also means "put a wedge in wood" in verbal usage. Kosarek Yale colexifies 'tooth' with 'name,' Yir Yoront with 'upwards,' Abzakh Adyghe *ce* also means 'feather, fur,' 'tooth,' and 'defense,' and Basque *hortz* also means 'prong,' 'pick' and 'hook,' as well as 'share.' Greek colexifies 'tooth' with 'tine' and 'bit sticking out,' Japanese with 'leaf' (the meanings being distinct prosodically), and Khalkha with "the graph ... in old Khalkha script." Haida *ts'ing* also means 'pestle,' Lakhota *hí* also 'to arrive, reach, get here, to come,' and Lesser Antillean Creole French *dan* also 'prong.' The Pawnee root *aar* also means 'to become, to do,' and Bororo *o* also 'front.' Aguaruna *dái* is also the name of a bird species. Miskito *napa* also means 'harrow, rake,' 'needle of a sewing machine,' 'claw, fingernail,' and denotes the "tooth" of a zip. Piro colexifies 'barb' and 'fishhook,' and for Tehuelche *?or ~ ?ol ~ or ~ o:r*, note the similarity with *?or ~ ?ol ~ or* 'nose, beak, point of knife,' there is another similar sounding word meaning 'perhaps.' Toba colexifies 'tooth' with 'placenta.' Hani *seq* also means 'to lead, guide, take along' and 'to marry a woman.' Hawaiian colexifies 'tooth' with 'beak of octopus,' 'nipper of an insect,' 'Aristotle's lantern,' 'interlocking stones,' and patterns on a mat or tapa. A Samoan term for 'tooth' can figuratively also refer to an 'enemy,' and another term colexifies 'goods, supplies,' 'fortune, wealth,' and 'business.' The meaning 'tooth' for this term is restricted to polite usage. Manange *1sa* also means 'land, soil, ground, mortar,' while Sedang *hónéng* appears to be derived from *héng* 'to have a toothache' by means of the nominalizing infix <ón>. Bislama colexifies 'tooth' with 'sucker, shoot' of a plant.

#### 145. *The Urine*

Representation: 74%

Motivated: 12.0%

Thereof Analyzable: 5.7%

Thereof Colexifying: 7.7%

Thereof by Contiguity: 0%

Thereof by Similarity: 2.7%

Most Commonly Associated Meaning: bladder, faeces, water, semen, gall, dung

Ngaanyatjarra, Tuscarora, Aguaruna, Huambisa, Yanomámi, and Bislama colexify 'urine' and 'bladder' (Anggor colexifies 'bladder' with 'to urinate' rather than 'urine'), while Ngambay has *kán-sín* 'bladder/gall, bladder-excrement.' Dongolese Nubian and Copainalá Zoque colexify 'urine' with 'faeces' as well as 'dung,' which latter association is also found in Indo-European, particularly Celtic (Buck 1949: 274). The association is realized formally

by alternation of nominal designants in Khoekhoe. Similarly, the Kashaya term contains an element meaning ‘to excrete,’ and there is a semianalyzable term for ‘urine’ featuring an element meaning ‘faeces’ in Sko.

Ngaanyatjarra and Fijian colexify ‘urine’ with ‘gall’ (Ngaanyatjarra also with ‘gall bladder’). Hawaiian has *hana-wai* ‘work-water,’ which also colexifies ‘irrigation’ and ‘menstruation,’ Welsh *dŵr* (*o’r bledren*) ‘water (of bladder),’ and Kiowa *sh’-tsoue* ‘burst-water’ (though the connection with the verb is somewhat unclear; the connection with ‘water,’ on the other hand, is also attested, though weakly, in Indo-European, evidenced by cognates in Ancient Greek and Sanskrit with the respective meanings, Buck 1949: 273). Semianalyzable terms with a constituent meaning ‘water’ or ‘liquid’ are found moreover in Kiliwa, Cayapa, Rama, and Tsafiki; there are also redundant terms featuring a verb meaning ‘to urinate’ as the second constituent. Tuscarora has the term *u?nhęhsú’kreh*, underlyingly /u-?nhęhs-ukr-eh/ ‘NOUN.PREFIX-egg/testicle-rubbish-NOUN.SUFFIX,’ for “foul or disgusting urine; a slovenly or slatternly person so fithly as to emit an odor of urine,” the stem of which also yields the term for ‘semen.’ Khoekhoe colexifies ‘semen’ with ‘urine’ by an archaic term, and Kwoma has *moku sobo* ‘semen raw/unripe/pure’ (*sobo* colexifies also other meanings).

Other associations include: Efik *i’kim* can also refer to “[a]ny disease of the urinary organs,” Hausa *k’ura*, meaning ‘urine’ and “drinking water in which nothing has been admixed” in the dialect of Sokoto also means ‘dust,’ and *fitsari* also to ‘humiliate a person in public.’ Yoruba has *ì-tò* ‘NMLZ-discharge.liquid,’ Baruya *suya* also means ‘tail,’ and Buin *iu* also ‘honey’ (the word is also a toponym for a mountain and a river). Burarra *darrjala* (which is also the name of a particular tree) might be related to *darrja* “be hot, as the sun shines hot,” and Kyaka *puu* also denotes the “Calamus or lawyer vine” and a ‘thread, rope’ among other meanings. Yir Yoront *kachl* also means ‘gill of fish,’ Haida colexifies ‘stale urine’ with ‘urine vessel, peepot’ and ‘amniotic fluid,’ and Pawnee *asuris* is also used with reference to the ‘odor of urine.’ Central Yup’ik *teq’uq ~ etquq* might be analyzable as /teq-quq/ ‘anus/bottom-one.like.’ Arabela *shaaca* also means ‘insipid, flavorless,’ Aymara has a term for ‘fermented urine’ which in fact also means ‘fermented,’ while Bororo *iku-ru* seems to be analyzable as ‘line-fire.’ Guaraní *ty* also means ‘juice’ and ‘big pile of something,’ and Jarawara *yoka* also means ‘athlete’s foot.’ Ancash Quechua *ishpay* also means ‘trench of large trees on small plants,’ and *pichi* is also an indefinite pronoun. Toba *lte* also denotes ‘impurities in water’ and ‘characteristic color, characteristic sign, mark.’ Bwe Karen *fi* also means “to be small, little; younger” and ‘to dazzle,’ Fijian *mi* also “to run in a small stream,” and Hawaiian *mī* also means ‘to dream’ inter alia. Finally, it should be noted that Haida distinguishes lexically between ‘fresh urine’ and ‘stale urine,’ and Wappo between urine from males and females.

#### 146. *The Uvula*

Representation: 27%

Motivated: 60.8%

Thereof Analyzable: 45.8%

Thereof by Contiguity: 21.3%

Thereof Colexifying: 15.0%

Thereof by Similarity: 38.8%

Recurrent associated meanings: tongue, throat, child/son, vagina/clitoris, tonsil, little, neck

The 'uvula' is frequently designated by complex terms of the lexical type constituting a metaphorical transfer from other body-parts or entities, with 'tongue' acting as the contiguity anchor. Among those, one pattern recurs cross-linguistically, namely that from 'child' or 'son' more specifically, as in Khalkha *keyken kele* /keyken kele(n)/ 'child tongue.' Such a term is also featured in Chickasaw, and the association is by colexification in Xicotepec de Juárez Totonac (the term also colexifies 'brood of animals,' 'loop of a net,' and 'interest'). Note also Tetun *nanarak-oan* 'palate-son.' Other complex terms are Khoekhoe *#khari-nam-s* 'small-tongue-3SG.FEM,' Badaga *kiru na:lange* 'little tongue,' Sora *akantalaŋən* /ə-'kanta:-'la'ŋ-ən/ 'POSS-branch-tongue-N.SFX,' and *le:r'laŋən* /l'e:r-'la'ŋ-ən/ 'grow.in.size-tongue-N.SFX,' Welsh *tafod bach* 'tongue took/hinge' and *tafod-ig* 'tongue-DIM,' Central Yup'ik *alungutayaaq* /alungun-taq-ya(g)aq/ 'tongue-device.for-little,' Guaraní *ape-kû-guy* 'skin-tongue-behind' (there is also the alternative term *ape-kuatî* 'skin-squirrel'), Yanomámi *akathamiki*, consisting of *aka* 'tongue,' *thami* 'internal part of body part' and the quantal classifier (see § 4.4.1) *ki* (this term colexifies 'velum' and 'gill'), Fijian *yame-leka* 'tongue-short,' Hani *lalngavq*, presumably /lalma-ngavq/ 'tongue-get.stuck,' and Yay *lin<sup>6</sup> kay<sup>2</sup>* 'tongue thing.' Finally, Itzaj has *ak'kal* 'tongue neck/throat' alongside *t'uy u-kal* 'vagina 3SG.POSS-neck/throat,' and in fact, this is another major recurrent association. Terms involving constituents meaning 'throat' and 'vagina' or 'clitoris' more specifically are also found in Carrier and Kiowa, and the association is by colexification in Wintu. 'Throat' and 'uvula' are colexified in Muna and Tuscarora; there is a semianalyzable term involving 'throat' in San Mateo del Mar Huave. Nez Perce has *múʔs-ries* /múʔs-eʔs/ 'swallow-INTR,' and there are semianalyzable terms betraying this association in Wintu and Piro. Finally, 'uvula' is colexified with 'tonsil' in Blackfoot, Badaga, and Lesser Antillean Creole French.

Other associations include: Hausa *beli* also denotes the "[a] small bud-like growth at the joints of corn-stalks," and is the name of various diseases inter alia. *Haki-n wuya*, another Hausa term, is analyzable as 'grass-GEN neck.' There is also another semianalyzable term where the identifiable constituent means 'neck' which also denotes a 'retropharyngeal abscess.' Basque *aho-gingil* is analyzable as 'mouth-lobe.' Greek colexifies 'uvula' with 'grapes,' while Haida colexifies 'uvula' with 'esophagus.' Embera features a semianalyzable term involving a constituent meaning 'way,' and Samoan *alelo* is also a term for the "eyes of a snake or eel."

#### 147. *The Vein*

Representation: 80%

Motivated: 65.0%

Thereof Analyzable: 20.0%

Thereof Colexifying: 46.8%

Thereof by Contiguity: 5.5%

Thereof by Similarity: 17.9%

Recurrent associated meanings: tendon, nerve, blood, way/street, muscle,

root, thread/string/fibre, pulse, gristle, line, lode, liana, rope, fishing line

The ‘vein’ is frequently colexified with ‘tendon, sinew’ (see also section 141). This is the case in Bakueri, Buli, Efik, Hausa, Dongolese Nubian, Yoruba, Baruya, Berik, Buin, Dadibi, Gurindji, Kyaka, Muna, Ngaanyatjarra, Nunggubuyu, One, Toaripi, Sahu, Kosarek Yale (marked with a question mark in the source), Yir Yoront, Chukchi, Khalkha, Nivkh, Sora, Highland Chontal, Haida (for younger speakers only), Lakhota, Itzaj, Xicotepec de Juárez Totonac, Tuscarora (colexifying ‘plantain’), Arabela, Cashinahua, Cavineña, Cayapa (by a semianalyzable term containing a constituent meaning ‘line’), Chayahuita, Guaraní, Lengua, Miskito, Tehuelche, Toba, Yanomámi (by the analyzable term *mathôyâhi* /*matha-yâhi*/ ‘leg-flesh’), Bislama, Fijian, Great Andamanese, Hawaiian, Lenakel, Malagasy, White Hmong, Rotuman, Takia, and Yay. Furthermore, Abzakh Adyghe has *l̥əntʃe*, containing *l̥ə* ‘blood’ and *ʃe* which perhaps bears the meaning ‘tendon,’ Miskito *tala wayka* ‘blood ligament,’ and there are semianalyzable terms featuring an element meaning ‘blood’ in Haida, Oneida, and Hani. Due to the association with ‘tendon,’ ‘vein’ is also colexified with ‘gristle’ in Welsh, Highland Chontal, and Haida, and with ‘line’ in Buli, Dongolese Nubian, and Bislama.

In general, associations are quite similar to those for ‘tendon, sinew.’ Buli, Hausa, Swahili, Toaripi, Yir Yoront, Basque, Chukchi, Khalkha, Sora, Itzaj, Lakhota, Xicotepec de Juárez Totonac, Arabela, Cavineña, Chayahuita, Embera (associated with different genders), Guaraní, Kaingang, Miskito, Tehuelche, Yanomámi (again, by the analyzable term *mathôyâhi* /*matha-yâhi*/ ‘leg-flesh’), Hawaiian, and Rotuman colexify ‘vein’ with ‘nerve.’ Muna, Ngaanyatjarra, Waris, Welsh, Guaraní, Lengua, Fijian, Hawaiian, Lenakel, and Tetun colexify ‘vein’ with ‘muscle’ (and Yanomámi with ‘muscular tissue’), and there might be a semianalyzable term on the basis of ‘muscle’ in Rama.

Rendille, One, Basque, Carib, Lengua, and Miskito colexify ‘vein’ with ‘root’ (similarly, Gurindji more specifically colexifies ‘single root of tree,’ and Hawaiian ‘small root, rootlet’); moreover, Kanuri has *zâr bú-bè* ‘root blood-of,’ and Bora *bájkymóóho* /*bájkyeémóóhou*/ ‘root-liana.’ Analogously, Copainalá Zoque colexifies ‘vein’ with ‘liana,’ and Tsafiki has *a’sán silí* ‘blood liana/string.’

Chukchi, Kildin Saami, Lengua, and Bislama colexify ‘vein’ and ‘thread, string’ or ‘fibre,’ while Yaqui has *ojbo wii’i* ‘blood thread.’

In fact, one difference between terms for ‘tendon’ and ‘vein’ is, as has emerged from the previous discussion, the frequent presence of ‘blood’ as a second constituent in complex terms for the latter acting as a contiguity anchor, as seen in Abzakh Adyghe and Miskito for the association with ‘tendon,’ in Kanuri for the association with ‘root,’ and in Yaqui for the association with ‘thread.’ A major difference with respect to ‘tendon,’ in spite of the many similarities, is also the presence of a transfer from ‘way, street’ to ‘vein,’ also with ‘blood’ acting as a contiguity anchor, as in Huambisa *numpa jinti* ‘blood way.’ This pattern is common in South America among the sampled languages, occurring alongside Huambisa in Aguaruna, Cavineña, Toba, and Wichí, but also attested in Mbum, Kolyma Yukaghir, Chickasaw, Santiago Mexquititlan Otomí, Malagasy, and Sedang. There are also some other complex metaphor-driven terms where one constituent is ‘blood.’ Hausa has *igiyar jinni* ‘rope blood,’ Ket analogously *sulaŋ* /*sül-àŋ*/ ‘blood rope’ (‘vein’ and ‘rope’ are colexified in Copainalá Zoque), Japanese has *kekkan*, perhaps analyzable as /*ketsu-kan*/

'blood-conduit,' Biloxi *haiti'* /hai-ti/ 'blood-house,' Kiliwa *khwat=h+yuul* 'blood=3+flow,' and Mandarin *xue4-guan3* 'blood-pipe.' Semianalyzable terms are found in Pawnee (which colexifies 'rubber') and San Lucas Quiaviní Zapotec.

Otherwise, Greek, Khalkha, and Embera colexify 'vein' with 'lode' (Greek figuratively also with 'talent'), and Lavukaleve and Bislama with 'fishing line.' Khalkha, Nez Perce, and Samoan colexify 'vein' with 'pulse,' and Vietnamese has *tinh mạch* 'calm pulse.'

Other associations include: Buli *jiin* is also the name of a "tightening string of a drum," Hausa *majinaciya* appears to contain *majina* 'mucus, snot' and also denotes a "weed with red fluid in it," Rendille *híy* also means 'sour milk' and 'relatives, kin,' while Baruya *wirila* also denotes "ridges on a shield" and "stingy fibres of wild yam." Dadibi *kigibili* is presumably analyzable as /kigi-bilibo/ 'maggot-walk.' Muna *ue* is also a unit of measurement inter alia, and Ngaanyatjarra *pulyku* (Northern dialect) also denotes the "Pencil Yam." Badaga *kuḍi* also means 'sprout, shoot' and 'penis' inter alia, Basque *zain* also "core, gist, crux" inter alia, Khalkha *sudal* also 'rings of a tree,' and 'ridge, stripe' inter alia, *xujang* also "rheumatitis, arthritis," Welsh *gwythien* also 'seam,' Ineseño Chumash also 'bow-string,' Kiowa *k'ih* also 'porcupine,' 'fire,' and 'to be heavy,' Lake Miwok *čúkem* also 'to be striped,' and Itzaj *wich'el* contains *wich'*, which can also mean 'wiry, thin.' Lesser Antillean Creole French *venn* also means 'seam.' Cubeo *pūpume* also means 'spiderweb' and the root *pūpu* indeed yields the meaning 'spider' when suffixed with a different classifier. Guaraní *tajygue* also means 'strength' and 'lungs,' and Piro *kotsa* also 'leaf' and 'child who is thin and underdeveloped.' Bwe Karen colexifies 'snake,' Hawaiian *a'a* may figuratively also refer to 'womb, offspring' as well as "to send greetings of love, joyous hospitality; joy at greeting a loved one," and Lenakel *nouanul-*, containing *noua-* 'mouth, opening, fruit,' may also refer to a "genealogical line" and a 'family, descent-group.' Tetun *uat* also is used with the meaning 'rain of wood,' Bislama *string* also means 'erection, to have an erection,' and Fijian *ua* is also used with the meaning 'wave, tide.' Finally, Yay colexifies 'tendon' and 'vein' with "violin, fiddle."

#### 148. *The Womb*

Representation: 64%

Motivated: 67.6%

Thereof Analyzable: 48.6%

Thereof Colexifying: 23.3%

Thereof by Contiguity: 22.8%

Thereof by Similarity: 29.8%

Recurrent associated meanings: child/son/daughter/baby, stomach, house, placenta/afterbirth, pregnancy, bag, place, netbag, guts, embryo/foetus, nest, receptable, sit, heart, vagina/vulva, bladder, basket, palace, mother, give birth

Terms for 'womb' (or 'uterus') are frequently metaphor-driven, with 'child,' or more specifically 'son' or 'daughter,' acting as the contiguity anchor. Terms with 'house' as the source concept, as in Yaqui *asoa-kari* 'daughter-house,' are found in Mbum, Ngambay, Yoruba (where 'embryo' rather than 'child' is the contiguity anchor), San Mateo del Mar Huave (where 'house' is colexified with 'nest,' note also Berik *tane gol* 'child nest,' that Bezhta colexifies 'womb' with 'nest' alongside 'bear,' and that Cashinahua colexifies

‘womb’ with ‘hole’), Cavineña, and White Hmong (see also Matisoff 2008: 176 for Tibeto-Burman specifically). Similarly, Japanese has *shi-kyū* ‘child-palace’ and Mandarin *zǐ-gōng* ‘child palace,’ the Japanese term being calqued. Alternatively, Central Yup’ik has *enliaq* ~ *neliaq* /*ena<sup>ə</sup>-liaq*/ ‘house-made,’ and Kyaka *romba anda* ‘stomach house/nest’ alongside *ingi anda* ‘intestines house/nest’ and *wane yakera petenge anda* ‘child embryo sitting house/nest’ (note that Kyaka *anda* also colexifies ‘nest,’ and hence this association is also present here). The association with ‘stomach’ and ‘child’ as contiguity anchor recurs in Baruya (*bwaramina* /*bwaranya-munya*/ ‘baby-stomach’) and Vietnamese (*dạ con* ‘stomach child’); very similarly, Mbum has *gûn-bîl* ‘child-abdomen’ (see also Hilpert 2007 for this association). Furthermore, Dongolese Nubian has *kumátté-n-tu* ‘vulva-GEN-stomach/interior’ (‘womb’ and ‘vagina’ are colexified in Carib), Meyah *ojóna otkonú* ‘married.woman stomach,’ and Ket *ām̄d hūj*, which contains elements meaning ‘mother’ and ‘stomach’ (compare also complex terms in Indo-European with a constituent meaning ‘mother’ reported by Buck 1949: 255, as well as Bislama *basket blong mama* ‘basket POSS mother/pregnant,’ and, in turn with the Bislama term, compare Fijian *kato-ni-gone* ‘basket/box-POSS-child’ and the colexification of ‘womb’ and ‘basket’ in Lenakel). ‘Womb’ is colexified directly with ‘stomach’ and/or ‘belly’ in Katcha, Yoruba, Kwoma (colexifying also ‘chest’), Toaripi, Badaga, Khalkha, Welsh, Carrier, Lesser Antillean Creole French, Nez Perce, Yaqui, Bislama (also with ‘pregnancy’ and ‘be pregnant’ in Bislama and one of the Khalkha terms), Hawaiian, and Rotuman (see also Buck 1949: 255 for evidence connecting ‘womb’ and ‘stomach’ in Indo-European). Moreover, Efik colexifies ‘belly of animal’ specifically (by a term that may be derived from a verb meaning ‘to conceal’), and in Miskito the association is present, but rare.

Returning to complex terms with ‘child,’ the Kyaka association with the notion of ‘sitting’ is also present in Kiliwa (*mnyis=waa-u* ‘foetus=sit-OBL,’ note that Mandarin and Khalkha colexify ‘foetus’ and ‘womb’ directly and compare the Yoruba term with a constituent meaning ‘embryo’ mentioned above). Khoekhoe has *fnû-lgau-s* ‘sit-be.left.over-3SG.FEM.’ In Badaga, a term for ‘womb’ is *ku:su pae* ‘baby bag.’ Similar terms on the basis of ‘bag’ as the source concept are also found in Kiowa and Piro, while Swahili has *fuko la uzazi* ‘bag of birth’ (alongside *chupa ya uzazi* ‘bottle of birth;’ a term based on a verb meaning ‘to give birth’ is also featured in Hausa and Toba). Hawaiian colexifies ‘womb’ with ‘bag,’ and there is a semianalyzable term in Highland Chontal (relatedly, Kildin Saami *pūŋŋ* is probably a loanword from Norwegian *pung* ‘bag’). An alternative Badaga term is *gabba pae* ‘pregnancy/happy.event bag,’ and ‘womb’ is colexified with ‘pregnant, pregnancy’ in Baruya, Khalkha, Bislama, and Rotuman. Piro has *whenewlu mapa*, containing *whene* ‘child’ and *mapa*, which colexifies ‘bag’ and ‘bladder;’ ‘bladder’ specifically is colexified with ‘womb’ in Hawaiian. Sora has *darakku:’onən* /*darakku:-’o:n-ən*/ ‘vessel/receptacle-child-N.SFX.’ Such terms are also found in Kashaya and Hawaiian, and the association is mirrored only somewhat differently in Toba (*l-co’o-oxo-qui* ‘3SG.POSS-give.birth-NMLZRECEPTABLE/ENCLOSED.SPACE’). A semianalyzable term where the identifiable constituent is ‘receptable’ is also found in Yanomámi. Moreover, Efik has *ēbiēt ēyën* ‘place child,’ and such terms are also found in Kanuri, Basque, Arabela, and Tetun.

There is also a wealth of complex terms with one constituent meaning 'child,' but the semantics of the other constituent being not among those discussed so far. Kolyma Yukaghir has *uon-könme* 'child-friend,' Chickasaw *oshaatoba*, analyzable as /oshi' aa-toba-'/ 'son DAT-become-NMLZ' and *oshq'to*, analyzable as /oshi' ǵlhto-'/ 'son be.inside-NMLZ' (this term is archaic), Haida *gid dl̥t'iisra*, containing *gid* 'child' and *t'iis* 'be in contact,' Lake Miwok *?élay šúkúhni*, containing *?élaj* 'child' and *šúkuh* 'stay, remain,' Oneida *-wilalahkwa?*, containing the roots *-wil-* 'baby' and *-l-* 'be in or on,' Santiago Mexquititlan Otomí *nt'oxubätsi*, which is analyzable as /nt'ots'i-bätsi/ 'granary-child,' Guaraní *memby-ruru* 'son-swollen,' Hupda *tǣ'h-yud* 'child-clothes' (marked with a question mark in the source), Kapingamarangi *mee dugu dama* 'thing put baby,' and Manange *1kola1tupΛ4k<sup>h</sup>ja*, containing *1kola* 'child,' *1tu* 'stay,' and *4k<sup>h</sup>ja* 'place.' Furthermore, Laz, Upper Chehalis, and Lake Miwok have derived terms from 'child,' and Badaga and Welsh colexify 'womb' with 'heart,' with the additional meaning being obsolete in Welsh.

Associations exclusively realized by colexification in the languages of the sample are: Hausa, Khoekhoe, Ngaanyatjarra, Sko, Khalkha, and Cahuilla colexify 'womb' with 'placenta' or 'afterbirth' (Hausa also with 'parents' and 'birthplace,' and Khalkha also with 'place to lie down, cave, den, lair'), Sahu, Badaga, and Sedang with 'guts,' and Buin, Burarra, Kwoma, and Takia with 'netbag' (Kwoma also with 'hill' and 'mountain,' inter alia).

Other associations include: Hausa *mahaifa* also means 'parents,' while Khoekhoe *llhās* also means 'ravine, gorge' and 'gully,' and Muna *tie* also "litter, time of giving birth (of animals)." Badaga *karu* also means 'heart' (in quotation marks in the source), 'dysentery,' "something that came from the stomach" and other things, and Khalkha *xabisu(n)*, apparently derived from *xabi* 'vicinity, neighborhood,' also means 'rib' and "trimming or metal plates on the bottom of a coat or mail." Nez Perce *?ilú't* may also refer to the "side, the part over the ribs of animals," and the Nuuchahnulth term *tiičšyāapi* is analyzable as /tīčš-yi-api/ 'alive-thing-stand' and colexifies 'life,' 'life principle' and 'childbirth.' Pipil (Cuisnahuat dialect) *xina:ch* also means 'ovary' and 'egg in chicken.' Tuscaraora *yēt?nēht?áhstha?* contains *-nēht-* 'to bury' (this term is marked as being unclear in the consulted source), Lengua *tathnak* might contain *tathna* 'navel,' Miskito *plauya* may rarely also refer to the 'bladder,' and Great Andamanese *ôtarain* seems to be derived from *ârain* 'gurjon tree' (*Dipterocarpus* sp.). Hawaiian *pū'ao* also means 'mesh of mats,' *ōpū* also 'tendon, vein, muscle' (the meaning 'womb' is said to be figurative in the source), 'crop of bird,' and 'disposition,' and *pupu'u* "to double up, draw the limbs together," hence 'foetal position,' and hence also 'womb' inter alia (this term may be derived from *pu'u* with the basic meaning 'protuberance'). Sedang *klea* also denotes the "inner edge of bamboo or of kōmea square strip of rattan." Samoan *fa'a-autagata* is analyzable as 'CAUS-be.shorthanded,' and Sedang *xoa* also means 'chest.' Tetun *knotak* also can refer to the 'waist.'

149. *The Wrinkle*

Representation: 48%

Motivated: 22.9%

Thereof Analyzable: 8.1%

Thereof Polysemous: 14.8%

Thereof by Contiguity: 4.6%

Thereof by Similarity: 15.5%

Recurrent associated meanings: crease/fold/pleat, ripple, skin, crumple, wither

Recurrent associations for this meaning are very few. There are some complex terms of the lexical type, where one of the constituents is 'skin,' with varying semantics of the second element. Abzakh Adyghe has *s'e-λe-r* 'skin/surface-be.located-SUFFIX' which also denotes a 'line' generally, and Piro *tslaha-mta* 'screen/grate/bars-skin' for "fine wrinkles in skin (not from age)." Furthermore, Koyraboro Senni has *kuurukuuru*, reduplicated from *kuuru* 'skin.'

Khoekhoe, Greek, Khalkha, Itzaj, Central Yup'ik, and Lesser Antillean Creole French colexify 'wrinkle' with 'crease,' 'fold,' and/or 'pleat,' Khoekhoe and Rotuman with 'crumple,' as of clothes, Greek, Welsh, Nunggubuyu, Bororo, and Fijian with '(have) ripple(s)' (similarly, Hausa colexifies 'wrinkling in water after long continuance' as well as 'emaciation,' while in Fijian the meaning colexified is 'wrinkle on brow' more precisely). Kosarek Yale and Yay colexify 'wrinkle(d)' with 'withered' (Kosarek Yale also with 'to become loose,' said of the hide of a drum).

Other associations include: Efik *ufrä* also denotes a 'leaping, jump,' and figuratively, "wrinkling, corrugating" among other meanings. The root of the Khoekhoe term for 'wrinkle,' *llai ~ llairo ~ lhō*, also means 'to shrivel of wet paper' as a verb. Yoruba *ì-wunjò ~ ì-hunjò* is analyzable as 'NMLZ-to.wrinkle.or.shrink,' and Gurindji *wanyjarrng* contains *wanyja* 'wrinkled yam.' Kwoma colexifies 'root,' while Kyaka *moo* can also refer to a mark or indentation in sand or metal. Nunggubuyu *-narmanarma-* also means 'to be furrowed.' Tasmanian terms in all varieties except the Northern for which data are lacking are said to also mean 'wart,' 'scar,' and 'tail,' and Basque *zimur* is also used with the meanings "empty chestnut," "ungrateful, thankless," and "tight, mean, stingy." Ket *kuraybet* contains *kud* 'bend' and *bed* 'make.' The Itzaj term *otz'tik* may also be used to refer to 'foam' or 'froth,' Lesser Antillean Creole French *pli* also means 'more, most,' Nez Perce *yukú'myukum* also 'crumpled,' while the Tuscarora root *-θriʔr-* yields both terms meaning 'wrinkle' and 'snail,' depending on the noun suffix attached. Wintu colexifies 'wrinkled' with 'stomach' and perhaps 'honeycomb tripe.' Central Yup'ik *imegglug-* is analyzable as /imeg-rrluk-/ 'roll.up/fold.up-one.that.has.departed.from.its.natural.state-', *qelengllak* also means 'scar, kink,' and *qacu-* also 'to be loose' and 'to sag,' Guaraní *chaĩ* also means 'bad' and 'ugly,' Miskito colexifies 'to wrinkle' with 'to rumple,' Wayampi *kala* also means 'rough' inter alia, Bwe Karen *θitru* also 'crumpled,' Lenakel *ulikulik* is reduplicated from *ulik* 'tough' (as of meat and other food), and Hawaiian *minomino* is reduplicated from the base *mino* 'dimple, depression, dent' and also means 'messed' with reference to dresses.



150. *The Dawn*

Representation: 86%

Motivated: 51.0%

Thereof Analyzable: 33.8%

Thereof Colexifying: 17.1%

Thereof by Contiguity: 30.7%

Thereof by Similarity: 8.6%

Recurrent associated meanings: light/light up/bright, day, morning, sunrise, night, come/arrive, sun, twilight, clear, dusk, tomorrow, cock crow, dark, white/become white, morning prayer, greeting, eye, land, mouth

The most frequent association is, unsurprisingly, that with 'day,' occurring in a variety of subtypes. 'Day' or '(be) day' and 'dawn' are colexified directly in Baruya, Wayampi, and Hawaiian. As for complex terms, 'light,' 'light up,' or 'bright' is the most common meaning of the second constituent (a meaning also commonly associated with 'dawn' in Indo-European according to Buck 1949: 993), as in Ket *kūn* /kə'n-i'/ 'bright-day' or Carib *emamili*, derived from *emami* 'be light.' Analyzable terms of the lexical type as in Ket are also found in Buli, Bislama and Tetun, and derived terms similar to that in Carib also in Abipón, Bora, and Jarawara. There are also variants of this pattern: Ngambay has *tà lò àrè* 'mouth time/day bright/lit.up' (with *lò àrè* being a term for 'day' itself), Kyaka *yuu nombalo* 'day first.light.of.day,' dialectal Basque *argi-haste* 'light-beginning,' Khalkha *sira gere* 'yellow light' and *gegegere-*, which is an inchoative verb derived from *gegere(n)* ~ *gege(n)* with the basic meaning 'daylight, morning daylight,' Itzaj *saska'tal* /sas-kab'-tal/ 'bright-world-come' (and variants of this term, including *jatz'katal* /jätz'-kab'-tal/ 'whip-world-come'), Mandarin *chen2-xi1* 'morning-sunlight,' Hani *aoq-bia bia* 'sky-bright/shining RED,' White Hmong *kaj-ntug* 'bright-sky,' Tetun *rai-naroma* 'land-to.grow.light,' and Vietnamese *bình minh* 'flat bright.' Moreover, Baruya, Buin, Badaga, Itzaj, Yana, Ancash Quechua, Hawaiian, and White Hmong colexify 'dawn' with '(day)light, bright' directly (ignoring glosses like 'first light of day'), and there are semianalyzable terms in Chukchi and Kaingang. The Itzaj association with an arrival by virtue of its term featuring a constituent meaning 'to come' is mirrored in Biloxi, which has *na''pi hu-di'* 'day come-??,' and such a term is also featured in Meyah. Similarly, Rotuman colexifies 'dawn' with 'come, arrive' directly, and Tsafiki has *oránan* /ora-nan/ 'good-come.closer.' Some languages have terms of the lexical type with a verb meaning 'to clear' or a noun 'clearness' as the second constituent, as in Buli *vari-nyaantiri* 'day clearness/brightness' and Muna *rara kamentae* 'clear morning.' Such a term is also found in Kiliwa, and Blackfoot and Aymara have semianalyzable terms with 'to clear, clear' as the meaning of the identifiable constituent (the Blackfoot term also denotes the 'clearing of weather'). Similarly, Rama has *sabítingi ngulaik* containing *sabítingi* 'clearing,' and Bakueri colexifies 'to dawn' with 'to clean;' note also the redundant Guaraní term *ko'e-t* 'dawn-clear.' Yir Yoront has *larr-mel-ngonngorr* 'day/place-eye-yesterday,' and 'dawn' is colexified with 'eye' directly in Koyraboro Senni among other meanings. Mali has *kunēnggunēng*, reduplicated from *kunēngga* 'sun, day,' Miskito *yu baiw-an* 'day sparkle-PAST.PTCPL,' Wichí *fwala'ihlo*, which contains *fwala* 'day, sun' and the locative suffix *-lo* 'in front of,' and terms betraying an association with 'sun' without colexification of 'day' are Badaga *ottu huṭṭu* ~ *hottu huṭṭu* 'sun be.born/rise,' Maxakalí *māyōn xupep* 'sun arrive/leave'

and Manange *1tiŋi-2p<sup>h</sup>ja* ‘sun-rise,’ with a semianalyzable term in addition found in Aguaruna. There are also a number of other complex terms of the lexical type where one of the meanings is ‘day.’ Buli has *vayok* /*vari-yok*/ ‘day-night,’ Efik *ñkpö-usen* ‘event-day,’ Basque *egun-senti* ‘day-feel,’ Highland Chontal *egaŋwayda lidine jouba*, containing *lidine* ‘day’ and *jouba* ‘finish,’ Lesser Antillean Creole French *pwen di jou* ‘point/fist/hand of day,’ Itzaj *chun-k’in* ‘base/trunk/foot-day’ (which also means ‘daytime, daylight’), Tuscarora *nyawé?nut*, containing *-eT-* ‘day’ and *-aT-* ‘stand’ (alongside another similar term), Cavineña *hucaca japada-ma* ‘day far-NEG,’ and Hupda *wág hi-yáét* ‘day FACTITIVE-lie.down.’ Moreover, there are semianalyzable terms making reference to the ‘day’ in Buli, Kaluli, Sahu, Welsh, Nez Perce, Chayahuita, Embera, Piro, and Hawaiian. For the association with ‘night’ in the Buli and Kiliwa terms mentioned above, compare Buli *saliuk yok* ‘morning nighttime’ (which can also refer to the ‘early morning’), Japanese *yoake*, analyzable as /*yo-ake-Ø*/ ‘night-end-NR,’ Cheyenne *vóone-ohtsé* ‘all.night-go,’ Piro *hoyetsno-kawa* ‘night-period.of.time.following,’ and that there is a term for ‘dawn’ which is derived from that for ‘night’ in Imbabura Quechua that can also mean ‘early.’

There are some complex terms for ‘dawn’ where ‘land’ is the identifiable constituent: Pawnee has *huraahrtaruuwispar*, containing *huraar* ‘be land,’ *ta-* ‘suspended,’ *iriwis* ‘across,’ and *war* ‘walk,’ and Tetun *rai-mutin* ‘land-white’ (compare also Burarra *ngana gunangarlcha*, containing *ngana* ‘mouth’ and *ngarlcha* ‘become white,’ Tehuelche *ʔoren ʔašk'en*, containing *ʔore ~ ʔore ~ ʔore* ‘be white’ and the locative nominalizer *-k'en*, as well as terms in various Romance languages going back to Latin *albus* ‘white,’ Buck 1949: 993), *rai-mutin* ‘land-abundant,’ and *rai-naroma* ‘land-to.grow.light.’

As already seen from various examples discussed so far, ‘dawn’ is associated with ‘morning’ in a number of sampled languages. Khalkha has a derived term, Buli *saliuk yok* ‘morning nighttime,’ Mbum *rim-péle* ‘dark-morning,’ Swahili *weupe wa alfajiri* ‘witness of early.morning,’ Muna *rara kamentae* ‘clear morning,’ Cayapa *dishquepenene* (/dishquepenene/), likewise analyzable as ‘dark morning’ (compare also the already mentioned Kiliwa *tiy chip* ‘dark/night clear/sweep’ and that there is a semianalyzable term with ‘dark’ being the meaning of the identifiable constituent in Upper Chehalis and one featuring a verb meaning ‘for darkness to disappear’ in Cashinahua), Rama *tamas aik* ‘morning side’ and Fijian *mataka lailai* ‘morning small’ (this term, however, also means ‘early morning’ itself), Hawaiian *ehu kakahiaka* ‘dust morning,’ also figuratively denoting ‘youth’ and ‘a shower that clears quickly,’ and *moku ka pawa* ‘be.cut DET darkness.before.dawn,’ and Mandarin *chen2-xi1* ‘morning sunlight.’ The meanings ‘dawn’ and ‘(early) morning’ are colexified in Buli (both by the term mentioned above and another one), Katcha, Yoruba, Muna, Badaga, Sora (by the term *duŋroi ~ duŋroilen ~ duŋrojən*, presumably containing *duŋ-* ‘get out of’ and the continuative marker *-roi*), Cahuilla, Lake Miwok, Bororo, Cayapa, Embera, Macaguán, Toba (by the term *yo’oxoñi ~ yi’oxoñi*, containing *yi’oq* ‘visible’), and Fijian, and there are semianalyzable terms in Basque, Haida, Wintu, San Lucas Quiaviní Zapotec, Huambisa, and Fijian.

Gurindji, Badaga, Basque, Wintu, Aguaruna, Embera, Maxakalí, Hani, and Tetun colexify ‘dawn’ with ‘sunrise’ (an association present in some Indo-European languages as

well, Buck 1949: 993), sometimes by analyzable terms discussed either here or in section 155.

Moreover, Sora colexifies 'dawn' with 'cock-crow,' and Kwoma, more explicitly, has *apochoko wo nedii* 'cock crow time;' such an analyzable term is also featured in Piro. Hausa and Swahili colexify 'dawn' with 'morning prayer,' and Khoekhoe, Yoruba, Toaripi, and Samoan with 'twilight' (an association evidenced in Indo-European in various languages by a variety of structural types, Buck 1949: 993), Toaripi by a semianalyzable term containing an element meaning 'weather,' and Samoan by a term related to a verb meaning 'blaze, flare up' as said of a fire. Lake Miwok, Ancash Quechua, and Wayampi colexify 'dawn' with 'tomorrow' (Ancash Quechua also with 'to wake up' and Wayampi also with 'the whole night through, until dawn'). Southeastern Tasmanian, Toaripi, and Aguaruna colexify 'dawn' with 'dusk,' and in Efik and Buin, relevant terms are also greetings used in the morning. Similarly, Fijian *kida* also means "to go and salute a person on his arrival."

Other associations include: Welsh *gwawr* also means 'hue.' Muna *hawo-hawo rusa* is analyzable as 'return.home-RED deer' (since deer return to the woods at dawn), and *kowine* is also the name of a particular star or planet. Rotokas colexifies 'dawn' with 'pink' and Sentani *he-bo-* is analyzable as 'hang-knock/strike-.' Badaga *kari hakki ja:ma* is analyzable as 'black bird time,' Japanese *akebono* contains *ake* 'rise,' and Cheyenne *hosóvoamaehtsé* and *hosóvoamáno'e* contain *hosóvo* 'backward.' Kiliwa (*tiy*)-*x-u?*+*saw-y* is analyzable as '(dark/night)-CAUS-OBL+see-ATT,' while Wintu colexifies 'dawn' with "spreading." Bororo has *baa aregodu* 'village appear,' Guaraní colexifies 'dawn' with 'year,' and Miskito has *lalma pauan* 'east reddened' and *lalma kahbi bara*, likewise containing *lalma* 'east' and *bara* 'when.' Hawaiian *iao* also denotes the 'silversides' (a kind of small fish), the planet Jupiter as the morning star, and is a toponym for a site in West Maui, while *kaiao* also means 'to enlighten.' Lenakel colexifies 'dawn' and "a shine, something shining," and Malagasy *mangirandràtsy* is analyzable as /*mangirana-ràtsy*/ 'to.have.chinks.through.which.the.light.shines-bad.' Samoan *vave-ao* is analyzable as 'early-dayflare,' and *tafa o ata* as 'be.visible LOC shadow.'

#### 151. The Day

Representation: 94%

Motivated: 53.9%

Thereof Analyzable: 4.5%

Thereof Colexifying: 49.4%

Thereof by Contiguity: 29.4%

Thereof by Similarity: 0%

Recurrent associated meanings: sun, daylight/light, time, date, weather,

noon, world, dawn, place/site, night in counting, occasion, hot/heat,  
week, watch/clock, morning, soil, afternoon, rain, sky, cloud, life, epoch/era

'Day' is very frequently colexified with 'sun' in the languages of the sample. This association is found in Hausa, Mbum, Rendille, Buin, Gurindji, Kyaka, Mali, Ngaanyatjarra, Sko, Toaripi, Japanese, Sora, Kildin Saami, Cahuilla, Ineseño Chumash, Comanche, Itzaj, Kiliwa (where the relevant term also means 'tawny, sun-colored'), Lake Miwok, Pawnee, Quileute, Xicotepec de Juárez Totonac, Yana, Yaqui, San Lucas Quiaviní Zapotec, Copainalá Zoque,

Lengua, Miskito, Rama, Tehuelche, Wichí, Fijian, Great Andamanese, Hawaiian, Manange, White Hmong, and Sedang. Furthermore, Dadibi has *giliga-de* ‘sun-with’ (in some of the sampled languages, there is also an association with ‘moon’ by virtue of the fact of it being colexified with ‘sun,’ compare section 60), Bororo *meri-ji* ‘sun-DET,’ and Tetun *loro-n* ‘sun-INAL.POSS;’ there are semianalyzable terms in Kwoma, Wappo, Kaingang, and Takia. By colexification of ‘sun’ and ‘watch, clock’ (compare section 79), the latter meaning is also colexified with ‘day’ in a small subset of these languages, namely Toaripi and Comanche. In Xicotepec de Juárez Totonac, the relevant term for ‘sun, day’ is derived from an element meaning ‘warm,’ while in Miskito and Hawaiian ‘sun, day’ is colexified with ‘hot, heat’ directly.

In Cahuilla, ‘time’ is colexified with ‘sun’ and ‘day’ in a single term, and the colexification of ‘day’ and ‘time’ alone is also attested in Efik, Ngambay, Kwoma, Kyaka (also with “event, opportunity, chance”), Meyah, Rotokas, Kosarek Yale, Yir Yoront, Sora, Carrier, Abipón, Aguaruna, Cavineña, Embera, Guaraní, Tehuelche, Great Andamanese, Lenakel, and Malagasy. Similarly, Yoruba, Basque, Khalkha, Highland Chontal, San Mateo del Mar Huave, Embera, Hawaiian, Samoan, and Tetun colexify ‘day’ with ‘date’ (San Mateo del Mar Huave also with ‘name’). Furthermore, Jarawara has *yama wehe* ‘thing/place/time light’ (on *yama*, see § 6.4.3.15.), and Ngambay, Kyaka, and Yir Yoront colexify ‘day’ with ‘place, site’ and related meanings, such as ‘soil’ specifically in the latter two languages. Perhaps remotely similarly, Cubeo, Piro, Hawaiian, and Rotuman colexify ‘day’ with ‘world’ (Piro also with “time sphere” and ‘expanse,’ and Hawaiian also with ‘to regain consciousness’ inter alia).

As already suggested by the Jarawara term just mentioned, another recurrent association is that with ‘(day)light,’ occurring by an analyzable term next to Jarawara in Efik (*uwem-ëyu* ‘life-time/daylight;’ compare colexification of ‘day’ and ‘life’ in Cayapa), and by colexification in Baruya, Buin (in which the relevant term *rua* also has the meaning ‘door’ due to accidental homonymy arising from borrowing), Kwoma, Kyaka, Sahu, Sentani, Yir Yoront, Badaga, Ket, Lesser Antillean Creole French, Nuuchahnulth, Guaraní, Hupda, Kaingang, Miskito, Yanomámi, Fijian, Hawaiian, Rotuman, and Samoan. The association is attested by semantic shift in Indo-European, as evidenced by cognates in Old Persian meaning ‘day’ and ‘light’ in Avestan, Greek, and Latin (Buck 1949: 991).

There are quite many cases in which ‘day,’ presumably more specifically in the sense of ‘daytime,’ is colexified with temporally contiguous phases of the day. Baruya, Wayampi, and Hawaiian colexify ‘day’ with ‘dawn’ (Yanomámi has *mi haru* ‘face to.dawn’ for ‘daylight,’ and Central Yup’ik has *erneq* /erte-neq/ ‘dawn-thing.that.results.from’), Maxakalí with ‘morning,’ (and Aymara with the later hours of the morning specifically), Efik, Ngaanyatjarra, Sko, Wintu, and Yuki with ‘noon,’ Yoruba and Ket with ‘afternoon’ specifically, and Meyah, Aguaruna, Cubeo, Guaraní, and Malagasy with ‘epoch, era.’ The Yana, Yanomámi and Samoan terms also mean ‘night,’ but this meaning only occurs in counting time.

Yoruba, Kyaka, Haida, Itzaj, Oneida, Embera, Lenakel, and Malagasy colexify ‘day’ with ‘weather’ (Haida also with ‘air,’ and Embera also with ‘atmosphere, climate’). Sentani and Bezhta colexify ‘day’ with ‘rain,’ Guaraní and Mandarin with ‘sky,’ and Hawaiian and

Samoan with ‘cloud.’ Kyaka, Sora, and Lenakel colexify ‘day’ with ‘occasion.’ Rotokas and Khalkha colexify ‘day’ with ‘week’ (Rotokas also with ‘garden, work’ and Khalkha also with ‘planet’).

Other associations include: Buli *vari*, as a verb, means ‘to seize, take by force,’ Hausa *yini* also denotes the “working part of the day” and the “withering of crops ... during the daytime” inter alia. Ngambay colexifies ‘day’ with ‘hour,’ while Burarra *ngorrngurra* also means ‘sleep’ and is also the common name for the gecko inter alia (since geckos are held against the skin to soothe children so they fall asleep). Kyaka *gii* colexifies ‘day’ with ‘season’ and ‘smile.’ The Muna term *gholeo* also means ‘dry season’ as well as ‘to dry in the sun’ when used verbally, while Nunggubuyu *arara* also means ‘to draw, to write.’ Abzakh Adyghe *mafe* contains *f(e)* ‘clear, bright,’ while Bora *cóóji#* also means ‘early,’ Wayampi *kōʔē ~ kōʔē* also ‘the whole night through, until dawn’ and ‘tomorrow,’ Hani *nao* also ‘to choke,’ Bwe Karen *mu* also ‘plant’ inter alia, and *ni* also “woman’s lower garment, skirt.” Lenakel *nian* also means ‘when,’ and among the meanings of Hawaiian *lā* are also ‘sail’ and ‘fin.’ Samoan colexifies ‘day’ with ‘celebration, party’ and ‘food for visitors.’

#### 152. The Dusk

Representation: 55%

Motivated: 55.6%

Thereof Analyzable: 31.6%

Thereof Colexifying: 24.8%

Thereof by Contiguity: 35.5%

Thereof by Similarity: 8.1%

Recurrent associated meanings: dark/darkness, twilight, sunset, afternoon, evening, night, sun, dawn, small, recognize, bad

Most frequently, ‘dusk’ is associated lexically with ‘dark,’ ‘darkness,’ or ‘to be dark.’ By colexification, this is the case in Buli, Kosarek Yale, Khalkha (where one of the relevant terms colexifies ‘clouded’), Welsh, and Lengua. As for complex terms, there is an interesting parallel between Lesser Antillean Creole French (*ti bwen* ‘small dark’) and Fijian (*buto-butō vaka-lailai* ‘RED-dark DERIV-small). Other complex terms betraying the association with ‘dark, darkness’ are Efik *ē’kīm ñk’pō usen’ubök* ‘darkness thing/event morning,’ Basque *ilunabar /ilun-nabar/* ‘darkness-gray,’ *ilun-alde* ‘darkness-side/region/area/proximity,’ and *ilun-senti* ‘darkness-feel,’ Upper Chehalis *kwetskwe’xta’n*, containing *kwets* ‘middle’ and *kwe’x* ‘dark,’ Chickasaw *okhlisht ishtaya* containing *okhlilish* ‘get dark’ and *aya* ‘go,’ Kiliwa *tiiy nip* ‘night/dark enter’ and *?-nyaay=kw+tiiy* ‘DN-sun/day=WH+dark/night,’ Guaraní *ka’aru-pytû* ‘afternoon-darkness,’ Bwe Karen *khi-la* ‘be.dark-down,’ and Tetun *rai-nakaras* ‘land-darken.’ There are semianalyzable terms in Buin, Aymara, Jarawara, and Kaingang, and note also the similarity between Nez Perce *siw’é’t* ‘dusk’ and *siw’é’t* ‘dark,’ as well as that between Comanche *tupisinawoni?* ‘dusk’ and *tupisibitū* ‘dark color.’ As the above discussion already makes clear, there are also associations with ‘night’ in some languages, sometimes due to colexification of ‘night’ and ‘dark’ (compare section 153). Alongside the associations in the abovementioned languages, and alongside direct colexification of ‘dusk’ and ‘night’ in Hani (by a semianalyzable term containing *aoq* ‘sky, heaven’), there are also many complex terms for the former where one of the constituents has the latter meaning. These are

Welsh *brig y nos*, containing *brig* ‘top, summit’ and *nos* ‘night,’ Highland Chontal *dihuama libuguih* /*dihuamna libuguih*/ ‘walk night,’ Quileute *?awí-sh* ‘night-become,’ Bororo *boeêo paru* ‘night beginning,’ and Chayahuita *tashi-rin* ‘night-CLASS.LARGE.AND.FLEXIBLE.’ There are semianalyzable terms in Baruya, Copainalá Zoque, Jarawara, Kaingang, and Tehuelche, and moreover, Biloxi has *pûspûsí’ ~ pû’spûs ~ pîspîsi* which seems to be reduplicated from *psi ~ pûs ~ pûsí* ‘night.’ The latter term colexifies ‘dusk’ with ‘twilight,’ an association also occurring (sometimes by one of the analyzable terms mentioned elsewhere in this chapter) in Efik, Khoekhoe, Rendille, Nunggubuyu, Toaripi (by a semianalyzable term featuring a constituent meaning ‘weather’ which also colexifies ‘early dawn,’ ‘dawn’ and ‘dusk’ are also not lexically distinguished in Southeastern Tasmanian and Aguaruna), Basque, Khalkha, Welsh, Wintu, and Hawaiian (by several terms, one of them being *mō-lehu* ‘QUAL/STAT-ashes/ash.colored’ which also means ‘tipsy’), while Badaga has *sande ja:ma* ‘twilight time’ (alongside another semianalyzable term containing *ja:ma*). There are associations with ‘(early) evening’ by colexification in Khoekhoe, Dongolese Nubian, Rendille, Upper Chehalis, and Quileute, and by analyzable terms in Yoruba (*werewere alẹ* ‘quick evening’), Central Yup’ik (*ataku-ar(aq)* ‘this.evening-little.piece.of’), and Samoan (*afiafi popogi* ‘afternoon/evening for.the.night.to.fall’), and by a semianalyzable term in Buin. In turn, Baruya, Kwoma, Nunggubuyu, Basque, and Bislama colexify ‘late afternoon’ with ‘evening,’ and, as in Samoan, there are complex terms in Kaluli (*ga:lo nudáb* /*ga:lo nudab*/ ‘afternoon taste’), Kwoma (*hogo ya yayi nedii* ‘late.afternoon sun ladder time’), and Guaraní (*ka’aru-pytû* ‘afternoon-darkness’), as already mentioned. The Kwoma term just mentioned colexifies ‘dusk’ with ‘sunset,’ and this is also the case in Dongolese Nubian, Badaga, Basque, Bora, Huambisa, Ancash Quechua (‘sunset with red color’ specifically), and Tetun. Internal structure of the terms, if present, is discussed in section 156. However, many of them contain an element meaning ‘sun.’ Such terms for ‘dusk’ are Kwoma *hogo ya yayi nedii* ‘late.afternoon sun ladder time,’ Kyaka *neta anda penge dokopa* ‘sun house departure when’ and *neta anda penge gii* ‘sun house departure time,’ Kiliwa *?-nyaay=kw+tiy* ‘DN-sun/day=PERF+dark/night,’ Wayampi *kwalai-ɔ-ʔa-ɔ-ɔ* ‘sun-3SG-fall-3SG-go’ (colexifying ‘west’), Hawaiian *li’u-lā* ‘slow-sun/day,’ Sedang *hài pɔxiammáng* ‘sun/day begin/commence,’ and Tetun *loro-teen* ‘sun-excrement,’ which also denotes a ‘species of moss.’ There are semianalyzable terms with an identifiable constituent meaning ‘sun’ in Kwoma, Cubeo and Huambisa.

Moreover, Kyaka has *yuu yasumi* ‘time last.light.of.day’ (and another semianalyzable term containing *yuu*) and Greek *likófōs* /*lík-o-fōs*/ ‘wolf-STEM.FORMATIVE-light.’ Interestingly, there are also three languages where terms for ‘dusk’ make reference to the fact that at dusk, it becomes hard to recognize the environment, and in particular other people. These languages are Muna (*dai wise* ‘bad face’), Rotokas (*vuri evei* ‘bad recognize’), and Sora (*erabmad’doja:lən* /*er-ab-mad-’doja’l-ən*/ ‘NEG-CAUS-recognize-relative-N.SFX’).

Although none of these meanings recurs exactly cross-linguistically, a number of the terms discussed above make reference to the fact that the sunset is a process by metaphorically using verbs of locomotion to convey the meaning, as in Chickasaw and Highland Chontal.

Other associations include: Buli *legi* also means ‘startled, shocked,’ Kyaka *yuu kwuange dokopa* contains *yuu*, meaning both ‘earth, ground’ and ‘day’ alongside still other meanings and *dokopa* ‘when.’ One *oi ninkle* is analyzable as ‘bed garden,’ and Abzakh Adyghe *pšepeze:χewey°e* as /pšape-ze:χewe-γ°e/ ‘horizon-expand-time.’ The Yuki term *hušámtik* contains *husám* ‘day before yesterday,’ and Cavineña *jaca-pude* is analyzable as ‘leave-to.color.’ Ancash Quechua *qarwayllu* appears to be related to *qarway* ‘to ripen, acquire color,’ while another Ancash Quechua term, *tsaqa*, also means ‘spine.’ Toba *napalñi* contains *apal* ‘be opaque.’ Tsafiki *pípuhuaqueno* contains *queno* ‘to do,’ and Great Andamanese *êrlōkorîtngrôî êr* ‘place.’ Kapingamarangi *heni* also means to “move around within an area,” Hawaiian *mōlelehu* also ‘drowsy, sleepy,’ and *li‘ulā* also means ‘mirage, hallucination.’

### 153. *The Night*

Representation: 96%

Motivated: 26.7%

Thereof Analyzable: 7.0%

Thereof Polysemous: 20.6%

Thereof by Contiguity: 24.6%

Thereof by Similarity: 0%

Recurrent associated meanings: dark/get dark/darkness, evening, day in counting, black, sleep, thing

The most common association for ‘night’ is that with ‘dark,’ ‘get dark,’ and/or ‘darkness’ (present also in Indo-European evidenced by cognates in Sanskrit, Old Persian, and Ancient Greek, Buck 1949: 992). In the sample, the association is realized by colexification in Anggor, Berik, Buin, Kyaka (where the relevant term also denotes a kind of tree), Ngaanyatjarra (also with ‘night sky’), Rotokas, Sahu, Badaga, Carrier, Upper Chehalis, Chickasaw, Highland Chontal, Kiliwa, Kiowa, Lesser Antillean Creole French, Yana (marked as dubious in the source), Yuki, Kaingang, Maxakalí, Miskito, Piro, Hawaiian, and Samoan. Ngambay has *lò-ndül* ‘time-black/dark,’ Burarra *ana-munya* ‘in-darkness,’ Dadibi *huli-de* ‘dark-who/with,’ Kyaka (*yuu*) *iminjingi dokopa* ‘(earth) darkness when,’ Bororo *boe êo* ‘thing dark,’ Jarawara *yama soki* ‘thing be.black/dark’ (compare § 6.4.3.15.), and White Hmong *tsaus-ntuj* ‘dark-sky,’ and derived terms, such as Ineseño Chumash *s-axiyi* ‘3SG/3SG.POSS-be.dark’ are found also in Nez Perce and Tuscarora; there is a semianalyzable term featuring a constituent meaning ‘dark, black’ in Lenakel. Furthermore, as for the association with ‘black’ in Ngambay and Jarawara, Buin, Kiowa, and Abipón colexify ‘night’ with ‘black.’

‘Night’ is colexified with ‘(end of) evening’ in Ngambay, Muna, Middle-Eastern and Southeastern Tasmanian, Abzakh Adyghe, Badaga, Welsh, Comanche, Lake Miwok, Wintu, Bwe Karen, and Yay (Yay also colexifies ‘late afternoon’), mirrored in a connection by the common Indo-European root for ‘night’ with ‘evening’ in Hittite (Buck 1949: 992). Two languages of Australia, Ngaanyatjarra and Nunggubuyu, by a typical example of actual/potential-polysemy (O’Grady 1960), colexify ‘night’ with ‘sleep’ (the relevant Nunggubuyu term also denotes the “ant lion larva”). Terms for ‘night’ are also used as a

unit of counting time (as ‘day’ is used in English) in Rendille, Buin, Nunggubuyu, Haida, Yana, Yanomámi, Hawaiian (where this usage is obsolete), and Samoan.

Other associations include: Buli *yok* also denotes a “rafter of a thatched house,” and Efik *okün'ëyu* ~ *okünoyu* appears to contain *ë'yu* ‘sunshine, daylight.’ Ngambay *tìl* also means ‘shrub,’ Buin *muu* also ‘other,’ Muna *alo* also ‘dew, expose to dew,’ Ngaanyatjarra *ngurra* also ‘home, camp site, house, bed, bedroom,’ and Sko *rángpang* contains *ráng* ‘sun.’ Toaripi *faita* is also the name of a type of red clay, and Yir Yoront *larr-ngonngorr* is analyzable as ‘day/time-yesterday,’ Badaga colexifies ‘night’ with ‘twilight,’ and is used adjectivally with the meaning “shadowy, shade” (another relevant term also means ‘to be, exist,’ and yet another one ‘to cover, cover up’). Basque colexifies ‘night’ with ‘to cover, close, shut, veil,’ while Japanese *yoru* also means ‘to twist’ (the meanings are distinguished in writing). The Cahuilla term *túkmiyat* ~ *túkmaat* contains *-túk-* ‘to go to bed, stay overnight.’ Wintu *cipi* colexifies ‘night’ alongside ‘evening’ also with ‘late,’ and Huambisa *suwe* also means ‘throat, chasm’ (original Spanish gloss is ‘garganta’). Bwe Karen *hε* also means ‘to wander about’ and ‘to be hot’ of food, and *nε* also denotes the spirit of a person. Hani *aoqqivq* also means ‘dusk’ (*aoq* is ‘sky, heaven’), Hawaiian *pō* also denotes the “realm of the gods,” “chaos, hell,” and ‘thick, dense,’ as said of flowers and fragrance. Kapingamarangi *boo* is also used with the meaning “feeling of foreboding” in compounds and also denotes a “woman who is pregnant for the first time.” Malagasy *àlina* is also a numeral for ‘ten thousand.’

#### 154. *The Noon*

Representation: 81%

Motivated: 65.8%

Thereof Analyzable: 58.3%

Thereof Colexifying: 9.2%

Thereof by Contiguity: 51.7%

Thereof by Similarity: 1.7%

Recurrent associated meanings: day, middle, sun, half, time, lunch, straight/right, afternoon, center, top, head, midnight, light/daylight, sky, hour, south, heart

Terms for the ‘noon’ are frequently analyzable of the lexical type, with one constituent being either ‘day’ or ‘sun,’ linked to the concept by contiguity. In a frequent type of terms, ‘middle’ is the meaning of the second constituent, as in Nuuchahnulth *?apwín n̄aas* ‘middle day’ and Meyah *mówa ot déis* ‘sun stand middle.’ These are also very common in Indo-European (Buck 1949: 996). In the present sample, such terms are featured alongside Meyah and Nuuchahnulth in Efik, Hausa, Kanuri, Ngambay, Swahili, Kaluli, Sahu, Basque, Greek, Khalkha, Kildin Saami, Welsh, Blackfoot, Ineseño Chumash, Lake Miwok, Lakshota (where in addition a verb meaning ‘to pass by’ is present), Santiago Mexquititlan Otomí, Tuscarora, Central Yup’ik, Copainalá Zoque, Bora, Bororo (where in addition an element meaning ‘sky’ is present, compare also Cavineña *barepatya* /barepa-patya/ ‘sky-middle’), Cubeo, Maxakalí (where the term is more specifically analyzable *mâyôn yăykote* ‘yūm /mâyôn yăykote yūm/ ‘sun in.the.middle.of sit/be.located’), Ancash Quechua, Bislama, White Hmong, and presumably Haida, although this is not ultimately clear from the consulted source. Similarly, Upper Chehalis has *Pôtûtsókûl*, containing constituents meaning



'middle' and 'set up.' Other similar terms, but without 'day' or 'sun' as contiguity anchor, are Buin *kugenu-page* 'part.of.roof-middle' (compare Huambisa *tutupnirmatai*, presumably /tutupnik-matai/ 'straight-ridge/crest' and Malagasy *mitatàovovònanana* (*ny àndro*) 'put.on.top ridgepole (?? day),' Cheyenne *-sétov-oéstá* 'middle-hang,' and Nez Perce *péq-pe* 'be.middle-LOC' (for 'at noon'). 'Noon' and 'middle' are colexified directly in Piro. In this language, as well as in Sahu and Welsh, 'middle' is colexified with 'center' (in Piro also with 'part, interior'), and this association is mirrored in Mbum (*làú sésèi* 'liver/center sun'); moreover, Bislama has *medel-dei* 'middle/center-day.' In Lake Miwok and Khalkha, 'middle' is colexified with 'half' (as well as with 'between' in Lake Miwok, and with 'midnight,' 'halfway, partly,' and 'mediocre, average' in Khalkha), and the terms in these languages therefore also betray an association with this meaning. They are not alone: Nivkh has *muyv-n'lami* 'day-half,' a structure which is also found in Basque, Welsh, Kashaya (where in addition a verb meaning 'to break into two' is perhaps present, though the overall analysis of the term is unsure; note also that Nunggubuyu colexifies 'dawn' with 'to be or become torn or split'), and Imbabura Quechua. This pattern is also common in Indo-European (Buck 1949: 996). Similarly, Ket *suran* possibly consists of constituents meaning 'half' and 'daylight,' and there are also terms in other languages making reference to '(day)light' specifically rather than 'day,' namely Efik (*uwem-ëyu* 'life-time/daylight') and Manange (*1t<sup>h</sup>on-4sol* 'light-bright;' this term colexifies 'noon' with 'afternoon'). Returning to terms with 'sun' or 'day' as contiguity anchor, in three sampled languages, Burarra, Toaripi, and Pawnee, there is an association with 'top,' either by terms containing elements with that meaning directly (Burarra and Toaripi; in the latter language, an additional element meaning 'head' is present), or by a verb meaning 'to stand on top' (Pawnee). Moreover, Yir Yoront has *par-thila* 'head/top-hole:ERG/LOC.' The association with 'head' is also borne out in Sora (*te'ŋbo:b'jəŋən* /te'ŋ-'b<sup>o</sup>:b-jəŋ-ən/ 'carry.on.head-head-sun-N.SFX'), Bororo (*barae etaiaidaia* ~ *brae etaia keje*, containing *barae* ~ *brae* 'non-Indian' and *aia* 'middle of head,' literally therefore "the time when the sun is on top of the head of the white people"), Lengua (*yitsikso ikhim* /yitsiksik ikhim/ 'crown.of.head sun/day'), and Miskito (by the term *won lal kat*, containing *lal* 'head' and *kat* 'until'). There are also complex terms, often featuring a constituent meaning 'day,' that make reference to the fact that noon is lunchtime, such as Kwoma *a yadii nedii* 'eat day time.' Bezhta has *łobał'as mex* 'noon.supper.gen time,' Kashaya *ma?a bumuyime?*, containing *ma?a* 'food' and *me?* 'time,' while Ngaanyatjarra, Chickasaw, Highland Chontal, Haida, and San Lucas Quiaviní Zapotec colexify 'noon' with 'lunch' (see also Buck 1949: 996 for this association in Indo-European). For Huambisa *etsa tutupin* 'sun straight,' which was mentioned above, there are also parallels. Yoruba has *òsán-gangan* 'afternoon-right,' Yaqui *luula-katek-o* 'straight-be.sitting-COND,' Rama *kíbing kíngík* 'straight day,' and Tsafiki *yotú* /yo-tu/ 'sun-straight.' Similarly, Berik has *gwere bolap*, containing *gwer* 'sun' and *bolap* 'summit.' Ngambay has *kàrè wùr énje* 'sun/time/hour heart mother,' and Takia *ad bibe-n* 'sun heart-3SG.'

There are also other complex terms with 'sun' or 'day' acting as the contiguity anchor. These are Yoruba *agbede-meji gjó* 'part-two day,' Muna *ghole-gholeo* 'RED-day,' Rotokas *ravireo vuuta* 'sun eventuate,' Toaripi *sare koko* 'sun/day narrow/contracted/restricted,' Itzaj *chumukk'in* /chu'm-Vk-k'in/ 'begin-DERIV-sun/day,' Kio-

wa *k'ih-sα* 'day-AUG,' Yana *baloorpa*, containing the verb *ba-* 'for the sun to move, be in position' ("sun is southward up the mountain"), Aguaruna *étsa tajímai*, a literal translation of which would be "the sun is between the two sides," Toba *yi na'aq*, perhaps 'richness day' (*yi* also has other meanings), Fijian *sigalevu tutu* /*sigalevu tūtū*/ 'day/sun-big stand,' Hawaiian *kau ka lā i ka lolo* 'place DET SUN AGT DET brains,' Hani *aoq-nao naohhaol* 'sky-day day-time,' Bwe Karen *lāmu chitha* 'sun be.overhead,' Malagasy *antoandro be* /*aN-to-àndro-be*/ 'LOC-exact/true-day-big/much,' Samoan *tū-tonu o le lā* 'stand-be.exact GEN DET sun' and *aoauli* 'day-be.pure,' as well as Tetun *loro-aas* 'sun-high' and *loro-natutu-n* 'sun-reach.its.peak-SINGULATIVE.' There is moreover a derived term in Carib, and Sko and Buli directly colexify 'sun' with 'noon,' while Efik, Ngaanyatjarra, Sko, Wintu, and Yuki colexify 'noon' and 'day.' There are semianalyzable terms where the identifiable constituent is 'sun' in Comanche, Aguaruna, Huambisa, and Rama, and where it is 'day' in Khoekhoe, Chukchi, San Mateo del Mar Huave, and Great Andamanese. Otherwise, the Cuisnahnuat dialect of Pipil has *wel-ō:rah* 'before/well-hour' and Cayapa *catyu'ura* /*catya-ura*/ 'high hour' (this term colexifies, as do Khalkha and Cahuilla, 'midnight;' furthermore, due to colexification of 'sun' with 'time' and 'hour,' these associations are also present in a complex Ngambay term with the other constituent meaning 'middle'). Oneida colexifies 'noon' with 'south' (Wayampi has *yane-alu-katu* 'our-south-good'), and Buli, Badaga, and Manange with '(early) afternoon.'

Other associations include: Abzakh Adyghe has *š(e)žay'e*, containing *ž(e)* 'to study, call out' and *y'e* 'time,' Badaga colexifies 'noon' with 'early morning' and 'off-hours' generally, while Khalkha *yde* is inter alia also the name of a river. Cahuilla *téklu-vel* is analyzable as 'stop-ABS,' and *máxeli-š* is derived from the verb *máxeli-* 'for the sun to be in the middle.' The Upper Chehalis term *čúq<sup>w</sup>=alm* is analyzable as 'set.up=erect.object,' while Hupda *hi-mĩ?-g'et* is analyzable as 'FACTITIVE-under-stand.'

### 155. *The Sunrise*

Representation: 52%

Motivated: 65.6%

Thereof Analyzable: 52.3%

Thereof Colexifying: 13.3%

Thereof by Contiguity: 16.5%

Thereof by Similarity: 40.7%

Recurrent associated meanings: sun, come out/go out/emerge, dawn, appear,

east, rise, light/daylight, day, sprout, arrive, split, be born, jump, burst out

Terms for the 'sunrise' are frequently morphologically complex and metaphor-based, with 'sun' (which sometimes is colexified with 'day') being the obvious contiguity anchor. Terms where the metaphor is based on notions such as 'to come out, go out, emerge,' such as Sora *duŋ'jəŋən* /*duŋ-jəŋ-ən*/ 'get.out.of-SUN-N.SFX,' are also found in Hausa, Burarra, Meyah, Kolyma Yukaghir, Biloxi, Chickasaw, Itzaj, Kashaya, Lake Miwok, Ancash Quechua, and Hawaiian. Similarly, Nez Perce has *tiñéhtit* /*tiñ-léht-it*/ 'sun/moon-out,' Kaingang *rã vỹ jur mũ ha* 'sun TOP coming go now,' and 'sunrise' is colexified with a verb meaning 'to emerge, come out' in Dongolese Nubian, Upper Chehalis and Rotuman; in the latter two languages the additional sense 'burst out' is present inter alia. Another common association is that with the meaning 'to appear' (which often also means 'to rise' in the sampled

languages), such as Chukchi *tirk-inini* 'sun-appear.' Such terms are also found in Khoekhoe, Muna (by colexification, also with 'to come up,' 'to break open'), Khalkha, Cheyenne, Hawaiian, and Tetun. Tetun also has the alternative term *loron-mosu* 'day-to.appear' (there are semianalyzable terms containing an element meaning 'day' in Kyaka, Highland Chontal, Comanche, Chayahuita and Huambisa). The Tetun term colexifies 'east,' a pattern of colexification also found in Yoruba, Burarra, Kolyma Yukaghir, Nez Perce, and Bororo, in all cases by analyzable terms discussed elsewhere in this paragraph. Wintu has the term *puy ʔel-kawal*, seemingly analyzable as 'east inside-dawn.' Yana has *zauduloozi*-, which is, according to the source, "evidently from *zaudulauxauzi* 'sun moves back out east-westward.'" 'Arrive' is also found in other languages which have complex terms of the lexical type as the meaning of the second constituent, as in Cubeo *aviá daino*, containing *aviá* 'sun' and *daino* 'arrival.' Such terms are also featured in Burarra, Kyaka (where an additional constituent meaning 'uphill, upwards' is present), and Maxakalí (where 'arrive' is actually colexified with 'leave'). Badaga has *ottu huṭṭu* ~ *hottu huṭṭu* 'sun/time be.born/rise,' and the 'birth'-metaphor is also attested in San Mateo del Mar Huave, which has *ajncnep teatnüt*, where *teatnüt* is 'sun' and *ajncnep* 'to be born, grow, bloom.' Alternatively, there is the term *ajnstop teatnüt*, with *ajnstop* meaning 'grow, bloom.' Similarly, Noni has *diuu ε san* 'sun to sprout,' and Khoekhoe *sore-s llhai-s* 'sun-3SG.FEM sprout/rise-3SG.FEM;' this association with 'sprouting' is also present in Hausa, where the verb meaning 'to come out' in the abovementioned term is colexified with 'to germinate,' as well as in Khalkha, also by colexification with another meaning, but in this case with 'to appear.' Kapingamarangi has *laa gu i hobo* 'sun DEC at jump,' and Samoan *oso a'e* 'jump go.up/climb.' There are also terms making reference to the 'brightness' and 'light' that the rise of the sun brings. Rama has *núnik áungai* /*núnik aungai*/ 'sun light,' Hani *aoq-bia bia* 'sky-bright/shining RED,' Tetun *rai-naroma* 'day-to.grow.light,' while Cashinahua and Cayapa colexify 'sunrise' with 'light' or related meanings, and there is a semianalyzable term where the identifiable constituent is 'bright' in Sáliba. Yoruba has *ilà oòrùn* 'split sun,' and an analogous term is found in Efik and Noni.

In several of the above discussed languages, such as Khoekhoe, the second constituent is alongside the primary meaning also glossed with 'to rise' when occurring in the context of sunrise. However, there are also some terms, such as Efik *utin asiakha* 'sun rise,' where the gloss of the complex term leaves it unclear whether the second constituent is only used in this context. Such cases are found also in Greek, Welsh, and Yuki. Genuine cases seem to be Nivkh *k'jeŋ myrfyr* 'sun place.where.something.rises time,' Fijian *ni cadra na siga*, where *cadra* also means 'to rise' of seeds and the younger generation inter alia, and Lenakel, which colexifies "reach or arrive at a place higher than from which one started" with 'to rise,' said of the sun and the stars.

Other complex terms where one of the constituents is 'sun' are Kaluli *kowo:na: of-o:* 'little.lizard sun-??,' Ngaanyatjarra *tjirtukarrany(pa)*, containing *tjirntu* 'sun, day' and *karra* 'twilight' (this term colexifies 'sunrise' with 'sunset'), Toaripi *sare patai* 'sun/day ascend,' Nuuchahnulth *hiisak\*istuuł?itq hupał* containing *k\*ist* 'move away' and *hupał* 'sun/moon,' Wappo *hìni čahwálse?* contains the word for 'sun, moon, clock, calendar' and a form of a verb meaning 'to crawl out.' Bororo has *meri ru-tu* 'sun fire-departure.' There are

semianalyzable terms where the identifiable constituent is ‘sun’ in Berik, Kaluli, Kyaka, Rotokas, Comanche, Guaraní, and Great Andamanese.

Gurindji, Badaga, Basque, Wintu, Aguaruna, Embera, Maxakalí, Hani, and Tetun colexify ‘sunrise’ with ‘dawn,’ Basque by the analyzable term *egun-senti* ‘day-feel,’ and Wintu by the analyzable term mentioned above.

Other associations include: Basque *goiz-argi*, also denoting the ‘light of dawn,’ is analyzable as ‘morning-light,’ and Upper Chehalis *pé.tekxwoihwo* contains the word for ‘night.’ Carrier *hae’aih en* is literally translated “where it comes from.” Nuuchahnulth *huupk<sup>w</sup>ista?at* is analyzable as /hup-k<sup>w</sup>isa-’ă’?a/ ‘roundish.spherical.or.chunky.object-move.away-on.the.rocks,’ while Pawnee *ta’a* is analyzable as /ta-a/ ‘suspended-come.’ Central Yup’ik *pit’e-* also means ‘to take game’ in the Yukon dialect, while Cayapa *dangueno* also means ‘to shine.’ Similarly, Tsafiki *chéino* and *cheyano* are related by unknown means to *chenu* ‘illuminate.’

#### 156. *The Sunset*

Representation: 55%

Motivated: 67.2%

Thereof Analyzable: 50.7%

Thereof Colexifying: 17.7%

Thereof by Contiguity: 14.2%

Thereof by Similarity: 45.1%

Recurrent associated meanings: sun, time, sink, enter, fall/drop, dusk, descend/go down/lower, day, disappear, die, west, twilight, dip, mountain, jump, sit down

As for ‘sunrise,’ of course, the meaning ‘sunset’ is expressed in many languages by complex terms where one of the constituents is ‘sun’ (which sometimes colexifies ‘day’). These are often metaphorical in nature. Among them, the most common subpattern is the association with terms meaning ‘to enter, entry,’ as in Khoekhoe *sore-s tǃǃâ-s* ‘sun-3SG.FEM enter/set-3SG.FEM.’ Such terms are also found in Ngambay, Yoruba, Comanche, Itzaj, and Cavineña, while in Nunggubuyu ‘to enter’ is colexified with ‘for the sun to set;’ the relevant term is in fact the reflexive of a verb meaning ‘to put into.’ Similarly, Rotokas has *ravireo rokore*, containing *ravireo* ‘sun’ and *roko* ‘die out, go into, penetrate into,’ Nez Perce *tiñeynékt* /tiñ-leyné-k-t/ ‘sun/moon-into,’ and San Mateo del Mar Huave *ajmel nüt* ‘enter day’ (for the association with ‘day,’ compare Kyaka *yuw kwualyamo*, containing *yuw* ‘day, light’ and *kwualyo* ‘wipe, brush away, spread, shake out’). Otherwise, associations with verbs meaning ‘to sink,’ as in Kolyma Yukaghir *jel’o:d’ad-amlujbâ* ‘sun-sink,’ are common (this term colexifies ‘west,’ which is also the case in the Bororo and Tetun complex terms already mentioned; Samoan has *tau-gāgaifo* ‘go-a.little.towards.the.west,’ and there is a semianalyzable term in Wintu which is also glossed as ‘red sunset’ and ‘cloudy and sun-shine’). Such complex terms also occur in Itzaj, Fijian, and Hawaiian, and the association is realized by colexification in Sora, Ineseño Chumash, Rotuman, and Samoan. Upper Chehalis *sʔísuʔs* contains *ʔísuʔ-* ‘to dive’ and a lexical affix ‘face, eye, round object, dollar.’ Meyah has *mówa esirí* ‘sun fall,’ and complex terms featuring verbs meaning ‘to fall, fall in’ or ‘to drop’ are also found in Hausa, Badaga (where an additional constituent meaning ‘time’ is present), Biloxi, Chickasaw, Itzaj, Bororo (where ‘fall’ is colexified with ‘be born’),

Ancash Quechua (where ‘to pounce’ is colexified with ‘to hide’ and ‘to bud’), and Tetun; moreover, the association is realized by direct colexification in Rendille, where the relevant term also means ‘go down’ and ‘attack.’ Cheyenne has *É-ta’ēhne* ‘3SG-disappear.’ Complex terms based on a verb meaning ‘to disappear’ are also featured in Ngambay and Wappo, while Ineseño Chumash and Central Yup’ik colexify ‘sunset’ with ‘disappear,’ the former language also with ‘to sink.’ Wappo colexifies ‘to disappear’ with ‘to die,’ note that in Muna, *soo* is alongside other things a euphemistic expression for ‘to die.’ The Wappo association with ‘die’ is also mirrored in Katcha by colexification. Complex terms on the basis of meanings such as ‘to descend,’ ‘to go down,’ ‘to lower,’ such as Kwoma *ya yeyi nedii* ‘sun go.down time’ are found in Efik (alongside an element colexifying ‘daylight’ and ‘time,’ the entire term colexifies ‘favor, privilege, opportunity’), Itzaj, Aguaruna, and Bislama (which also has *san i draon*, with *draon* meaning ‘to drown’ as well as ‘for the sun to set,’ inter alia); moreover, Buli and Rendille colexify the relevant meaning directly. Noni has *diuu ε sele* ‘sun to jump,’ and ‘sunset’ and ‘fly, jump’ are colexified in Lenakel. Wintu has *sasun p<sup>h</sup>uyuuq ʔolwiti harasin*, containing *sas* ‘sun, moon,’ *p<sup>h</sup>uyuuq* ‘mountain,’ *har* ‘move’ and *asin* ‘away,’ and a semianalyzable term containing a verb meaning ‘to move’ is attested in Yana (incidentally, ‘hill, mountain’ is among the meanings colexified with ‘for the sun to set’ in Rotuman). Kashaya has *ha’da cahci* ‘sun sit.down,’ and a term featuring a verb with the meaning ‘sit down’ alongside ‘sun’ is featured in Burarra as well. Greek and Welsh have terms containing elements meaning ‘sun’ and ‘set.’

‘Sunset’ and ‘dusk’ are colexified in Dongolese Nubian, Kwoma (by the analyzable term *hogo ya yayi nedii* ‘late.afternoon sun ladder time’), Badaga (by the analyzable term *sande ja:ma* ‘twilight time’), Basque (by the analyzable terms *ilun-alde* ‘darkness-side/region/area/proximity/nearness’ which itself colexifies ‘twilight’ and *ilunabar /ilun-nabar/* ‘darkness-gray’), Bora, Huambisa, Ancash Quechua (colexifying ‘sunset’ with ‘red color’ specifically), and Tetun (by a semianalyzable term containing *rai* ‘land’), often also by analyzable terms discussed elsewhere in this paragraph, and Arabela has *nio nininiutianiyani* containing *nininiu* ‘to be dusk,’ note also the similarity between Xicotepec de Juárez Totonac *smalanka’nan* ‘for the sun to set’ and *smalanka’n* ‘dusk’ and Miskito *sáiwawa* ‘sunset’ and *sáíwan* ‘dusk.’ Efik colexifies ‘sunset’ with ‘evening,’ and Dongolese Nubian with ‘early evening’ specifically.

Other complex terms featuring a constituent meaning ‘sun’ are Kyaka *netame anda pelyamo/penge* containing *neta* ‘sun,’ *anda* ‘house’ and *penge* ‘arrival’ (there is another term featuring *neta* and *anda* as constituents the structure of which is not clear otherwise), Ngaanyatjarra *tjirntukarrany(pa)*, containing *tjirntu* ‘sun, day’ and *karra* ‘twilight,’ Toaripi *sare sukopai /sare sukapai/* ‘sun dip.into’ (considered likely in the source; “[t]o the Elema, being a coastal people, the sun when setting seems to dip into the sea;” Kosarek Yale colexifies ‘to dip into water’ directly), Khalkha *nara(n) singge-ky* ‘sun absorb/set-??,’ Nivkh *k’jeŋ uy yr* ‘sun get.to mouth,’ Itzaj *b’el k’in* ‘sun go/travel/trip/going/exit,’ Lesser Antillean Creole French *sòlei kouché* ‘sun lie/sleep,’ Aguaruna *tsawáut /etsá-wáut/* ‘sun-open.mouth,’ Rama *núnik ausam altuang* containing *núnik* ‘sun’ and *altuang* ‘waiting,’ and Hani *naolma xavq li qavq* ‘sun stay.overnight go PAST.’ Further, there are semianalyzable terms where the identifiable constituent is ‘sun’ in Berik, Dadibi (where ‘sun’ and ‘fire’ are colexified),

Kaluli, Kyaka, Basque, Chukchi, Cubeo, Guaraní, Huambisa, Maxakalí, and Great Andamanese.

Other associations include: Ngaanyatjarra uses the same term for ‘sunset’ and ‘sunrise,’ Blackfoot *isttahkapi* also means ‘to crawl in, under,’ and *otahkoomaiksistoyi* contains *otahkoo* ‘orange, yellow.’ Carrier *na-e-aih en* is analyzable as ‘to.the.ground-??-get when,’ whereas Upper Chehalis *stáqnč* contains *táqa-* ‘cover/shade.’ Kwi’ *suenta’ne*, in contrast, contains *kwi*’s ‘get dark, night.’ Comanche *tabe?ikai ~ tabe?ikaru* also means ‘evening,’ while a literal translation of Kiliwa *ruwp ?ii* might be “shrouded from view.” Pawnee *astaruukita* contains *as-* ‘foot’ and *huukita* ‘be on top’ (“a metaphor, lit. the sun has its feet on (the horizon)”). Tuscarora colexifies ‘sunset’ with ‘starset,’ and Wintu has *puyel hololbe*, containing *puyel* ‘east hill’ and *holol* ‘sunshine, bright, light.’ Chayahuita *i’hua-raya* is analyzable as ‘a.short.time.ago-CLASS.FACE.EYE.OR.SEED.’

#### 157. *Man (Human Being)*

Representation: 83%

Motivated: 35.1%

Thereof Analyzable: 8.9%

Thereof Colexifying: 26.2%

Thereof by Contiguity: 7.2%

Thereof by Similarity: 0.0%

Recurrent associated meanings: member of ethnic group, man, body, thing, owner/proprietor, mankind, woman, husband, earth/ground, child

Apart from use of the relevant terms as indefinite pronouns (‘somebody’), which is disregarded here, terms with reference to ‘human being, person’ are commonly colexified with ‘man’ in the sense of ‘male person’ specifically, a pattern familiar from e.g. English (see Buck 1949: 79-80 for discussion of the broader Indo-European context). Precisely because of this, it is difficult to assess the strength of this pattern cross-linguistically, since the gloss ‘man’ is in fact ambiguous. Therefore, only cases are reported in which the glosses give reason to believe that ‘male’ specifically is genuinely colexified. This is the case in Efik, Kwoma, Yir Yoront, Basque, Chickasaw, Highland Chontal, San Lucas Quiaviní Zapotec, Cashinahua, Wayampi, and Bislama, while in Khoekhoe the meanings ‘human being’ and ‘man, husband’ are formed by using the same root, but suffixed with different nominal designants. Moreover, Sentani has the dvandva compound *do-mije* ‘man-woman,’ and Takia *tamol-pein ~ tal-pein* ‘man-woman’ (this type of compound is common in New Guinea in general; for a precise parallel in the term for ‘people’ in Kalam see Pawley 1993: 99). Such a term is also found in Xicotepec de Juárez Totonac. Furthermore, Kanuri has *k-âm*, containing the prefix *k-* and the plural form of ‘man,’ and Guaraní has *yvy-póra* ‘earth-dweller’ (compare also Kaingang *êprā ke* ‘on.ground make/say’ and Latin *homo*, related to *humus* ‘earth’).

Another common association is the colexification of ‘human, person’ with a member of one’s ethnic group in particular. This occurs in Koyraboro Senni, Burarra (by the term *gugaliya*, derived from *galiya ~ jaliya* ‘to hear’), Gurindji, Nunggubuyu, Yir Yoront, Nivkh, Cahuilla, Haida (also by the redundant terms *xàaydlāa xàaydaraay* ‘people visible.world/home,’ which refers to the Canadian Haida specifically, and *xàayda giits’aads*

'people servants'), Lake Miwok, Quileute, Wintu, Yuki, San Lucas Quiaviní Zapotec, Bororo, Guaraní, Hupda, Miskito, Piro, Tsafiki, Yanomámi, Hawaiian, and Samoan. Furthermore, Efik, Ngaanyatjarra, Sentani, Yir Yoront, Copainalá Zoque, and Miskito colexify 'human being, person' with 'body' (Miskito also with 'flesh'), Yoruba, Badaga, and Hawaiian with 'mankind,' Kwoma and Bislama with 'husband,' and Ngambay, Bislama, and Hawaiian with 'owner, proprietor.' In Khoekhoe, the meanings 'man' and 'husband' are derived from the same root with different nominal designants (shift from 'person, human being' to 'male' and 'husband' is attested in Slavic according to Buck 1949: 80).

Kiowa has a general term 'human, person' identical segmentally to a verb meaning 'be alive' which takes additional suffixes depending on whether a man or a woman is referred to; a semianalyzable term probably featuring a constituent with that meaning is also found in Kiliwa and in Khalkha, where it colexifies 'animal.' Efik, Bororo, and Hawaiian colexify 'human, person' with 'thing' (Bororo also with 'time' and Hawaiian also with 'to say' inter alia), and in Swahili, the same root, associated with different prefixes, conveys these meanings. Koyraboro Senni *ibuna'adamayze* contains *ize* 'child;' the first constituent is borrowed from Arabic *'ibn 'aadam* 'son of Adam' (Hausa *'dan adan*, which has a similar structure is, when used in the singular, often applied to a person who has done something wrong), and Sahu and Hawaiian colexify 'person' and 'child' inter alia. Moreover, there is a semianalyzable term featuring a constituent meaning 'son, child' inter alia in Hani.

Other associations include: the formally redundant Ngambay term *kèje lè dèw* is analyzable as 'think GEN person' (*dèw* also means 'soul;' note that the Indo-European root giving rise to Engl. *man* etc. is, on one interpretation, connected with *\*men-* 'to think,' Buck 1949: 80). Buin *roi* also means 'gallbladder,' and Berik *angtane* also 'passenger' specifically. Kwoma colexifies 'human, person' with 'adult, mature,' 'front,' 'top,' 'finger, toe,' and 'twenty.' Muna *mie* is also used with reference to 'animals that build nests,' while Ngaanyatjarra *yarnangu* can also refer to the 'whole of something' and the 'appearance of something,' and Sentani *u* also means 'empty,' as of things. Tasmanian (Middle-Eastern) *kekána* contains *kána* 'language, utterance, speaking.' The Chukchi term *?orawetl?an* contains *?oraj* 'openly' (the term is explained by Bogoraz 1922: 828 as quoted by Fortescue 2006: 269 as literally meaning 'one who walks openly,' which is used as a name for humans by malevolent spirits in Chukchi myth). Khalkha *kymyn* also means 'personality' and Lesser Antillean Creole French *moun* also means 'world, universe.' Blackfoot colexifies 'person' with 'pupil,' Itzaj *kristiyaanoj*, a loanword from Spanish, also denotes a 'Christian' specifically, and *mak ~ maak*, a native term, also means 'top.' Embera *ẽmberá* means 'person, personage' with masculine gender and 'people' with feminine gender. Guaraní *tekove* contains *teko* 'nature, character, being' and colexifies 'life' and other meanings. Huambisa *shuar* also means 'enemy,' and Hupda *húp* also 'good, new,' while Wayampi *te-kə* is analyzable as 'NON.DETERMINATION-to.be.' Fijian *tamata* can also refer to chicken still in their eggs, and Hawaiian *kanaka* also means 'population,' 'pregnant,' and 'inhabited' inter alia.

158. *The Saturday*

Representation: 59%

Motivated: 33.9%

Thereof Analyzable: 33.8%

Thereof Polysemous: 0.2%

Thereof by Contiguity: 22.9%

Thereof by Similarity: 4.5%

Recurrent associated meanings: day, six, week, Saturn, Sunday, small, work, prepare, younger sibling, unique

Terms for ‘Saturday’ are overwhelmingly complex and of the lexical type in the languages of the sample, with ‘day’ (which sometimes is colexified with ‘sun’) being the meaning of one of the constituents. Apart from semianalyzable terms in Basque, Chayahuita, Guaraní, Fijian, and Manange, most frequently, ‘six’ is the meaning of the other constituent, as in Sko *bang nápánghi* ‘day six’ (*bang* is not explicitly glossed in the source, but since it recurs in the names of all other days of the week, it seems safe to assume that its meaning is ‘day’). Such terms are also found in Ket, Chickasaw, Hani (where an additional constituent meaning ‘week’ is present; there is another semianalyzable term with this element), Hawaiian, Kapingamarangi, and Rotuman (where it is used by members of the Roman Catholic church). Similarly, Kaluli has *do:go:fe-yá* (/do:go:fe:-ya/) ‘six-right.here’ for ‘on Saturday,’ Mandarin *xíngqī-líu* ‘week-six’ and *lǐ-bái-líu* ‘ritual-worship-six’ (there are also other complex terms featuring a constituent with the meaning ‘week’: Swahili has *juma-mosi* ‘week first,’ Kashaya *capasi me?*, analyzable as /capa ši-w me?/ ‘week make-ABS time,’ ‘Saturday’ in Oneida is derived from a term meaning ‘daylight, week,’ and ‘Saturday’ and ‘week’ are colexified directly in Central Yup’ik by the analyzable term *maqi-neq* ‘steambath-thing.that.results.from’). Ngambay has *ndò kùlà mǐsán* ‘day work six,’ and Kiliwa *tí?chat=msirlhpaayp* ‘work=six’ (note also the rare Hawaiian term *lā ho’omalolo* ‘day cease.work’). In contrast, Vietnamese has *ngày thứ bảy* ‘day ordinal seven,’ the variation presumably due to differences in which day is the day of rest and prayer and differences in which day is taken to be the first of the week.

Japanese has *do-yō* ‘earth/Saturn-day’ and Welsh *dydd Sadwrn* ‘day Saturn.’ ‘Saturday’ is colexified with ‘Saturn’ in Badaga and Khalkha (see § 6.4.3.13.6. for the history of this pattern, as well as Buck 1949: 1007-1009 on the coexistence of the ‘ecclesiastical’ system and the ‘planetary’ system in Indo-European). Abzakh Adyghe and Basque have *mefezaq<sup>2</sup>e* and *egu-bakoitz* respectively, both analyzable as ‘day-unique.’

Rotokas has *Topekakau voki* ‘tobacco day’ (“comes from the old practice of paying wages with tobacco on Saturday;” compare also the fact that an Upper Chehalis term for ‘Saturday’ contains a word for ‘time’ and a verb meaning ‘to distribute,’ with the explanation in the source offered having recourse to the fact that on Saturday rations were distributed). A literal translation of Blackfoot *to’tohtáátoyiksistsiko* according to the consulted source is “the day before the holy day,” that is, Sunday, Chickasaw *Nittak Hollo’ Nakfish* contains *nittak* ‘day’ and *nakfish* ‘younger sibling of the same sex,’ Tuscarora has *awé?nakwt*, literally ‘day alongside,’ Fijian *Siga Vakarau leka*, containing *siga* ‘day’ and *leka* ‘small,’ Samoan *Aso To’ona’i* ‘day collect.food.in.preparation.for.Sunday.meal’ (Lenakel has (*nian taha*) *n-epinapine-aan* /(nian taha) *n-epinapina-aan*/ ‘(day BENEFACTIVE) NMLZ-prepare-



NMLZ,' a term which is restricted to the speech of older Christians, and Rotuman (*terán*) *a'itā* /(*terāni*) *a'itā*/ '(day) prepare;' this term is used by Wesleyan Methodists). All the denominations just mentioned, starting with that in Blackfoot, make reference in some way to the fact that 'Saturday' is contiguous temporally to 'Sunday,' the most important day of the week. The Fijian association with 'small' is explained by the fact that the 'Saturday' is a "small" holiday when compared with the Sunday, and this interpretation is corroborated by terms such as Cheyenne *Tšêške'-ma'heóneéšēva* 'little-Sunday,' Kiowa *dā'k'ih-syḥn-gyḥ* 'Sunday-small-NOUN.POSTFIX,' and the fact that Carrier has a semianalyzable term where the identifiable constituent is 'small.' Moreover, Biloxi has *noxwi'so'tka* /*noxwi'di so'tka'da*/ 'Sunday younger.brother,' and there is a semianalyzable terms involving a word for 'Sunday' in Haida ('Sunday' in Haida is *s@ndii*. Haida also features the term *sandiigaa cajuu*, where *cajuu* is 'small,' and *sandiigaa* is also suspiciously similar to Engl. *Sunday*. The source notes that this term is "[s]aid to derive from the missionary habit of flying a small flag on Saturday and a big one on Sunday," which suggests that this is a term of the Cheyenne and Kiowa type just mentioned) and Nez Perce (where the literal translation offered in the source is "toward Sundaying"). Moreover, Kyaka has *koro kuki* 'period.of.time small' (this term also denotes a 'pause').

Other associations include: Buli *Asibi* is also a name given to children born on a Saturday, and Badaga *cani* ~ *sani*, colexifying 'Saturn,' also means 'bad omen, ill fate' and similar things. Pawnee *Piiriku* contains *iirik* 'to see,' the literal translation and explanation given in the source being "seeing different ones (when people went into town on Saturdays to shop)."

#### 159. *The Virgin*

Representation: 38%

Motivated: 53.3%

Thereof Analyzable: 53.0%      Thereof Colexifying: 2.7%

Thereof by Contiguity: 36.0%      Thereof by Similarity: 7.1%

Recurrent associated meanings: girl/daughter, unmarried girl/woman, man, NEG, closed/shut, young, female/woman, husband

Terms for 'virgin' are in some languages of the sample analyzable, containing an element meaning 'man' or 'husband' and a negator alongside other morphemes. For one, Ngambay has *gère ngàw ááng* 'know man/husband NEG,' and similar terms, with verbs meaning 'to feel' and 'to see' rather than 'know' respectively are found in Chickasaw and Bora. Kyaka has *wanakeme akali nyii range*, containing *wanake* 'girl,' *akali* 'man, husband,' *nyii* 'take!' and *range* 'self, ego' (alongside other semianalyzable terms featuring a constituent meaning 'man, husband'), and Piro *makloji jeji metkatowa*, containing *makloji* 'girl,' *jeji* 'man,' and *metkatowa*, which seems to be related to *metkatu* 'blind.' Ngaanyatjarra *wati-ku ngurra* is analyzable as 'man-of ignorant.' Embera has *w'ëra aw'ërakirú*, containing *w'ëra* 'woman' and *aw'ërakirú*, an adjective meaning 'new, virgin.' Fijian *goneyalewa savasavā* is analyzable as 'girl clean,' and *dau lato* seems to contain *dau* 'to commit adultery' (*lato* is the name of a tree species).

There are derived term from roots meaning ‘unmarried’ in Khoekhoe and Lenakel (‘virgin’ is colexified with ‘unmarried girl/woman’ in Kwoma, Kyaka, Nez Perce, Pawnee, Central Yup’ik and Lenakel). Samoan colexifies ‘female’ generally with ‘virgin,’ semianalyzable terms where the identifiable constituent is ‘woman’ are featured in Kwoma and Guaraní, and Tetun has *feto-raan* ‘woman-bleed,’ referring to menstruation. This is also the case in Upper Chehalis (*máymáyšxamri* contains *máya-* ‘enter’ and *šam-* ‘menstrual period’). A term referring to the physical rather than cultural aspects of virginity is also found in Yanomámi (*ka kōmi* ‘opening closed’). Yir Yoront and Hawaiian have similar terms of the lexical type, and Efik a derived term from a verb meaning ‘be closed.’

Haida and Miskito colexify ‘(be) virgin’ with ‘(be) young’ (Japanese has the term *otome*, questionably analyzable as /oto-me/ ‘young-woman’), while in Muna, Khalkha, Welsh, Nez Perce, Pawnee, Samoan, relevant terms can also refer to a ‘(teenage) girl,’ ‘daughter,’ and sometimes even ‘child’ more generally.

Other associations include: Noni *wan wvu tfu* contains *wan* ‘child’ and *tfu* ‘each,’ Baruya *kwaiyagaala muja* makes reference to the fact that it is forbidden to address a virgin in the male speech (*yagaala*) concerning marriage, and Burarra (-)yawuk also denotes a ‘childless girl or woman.’ The Muna term *bungasa* also means ‘untouched,’ as said e.g. of palm trees not yet tapped and *kalambe*, as a verb, is glossed as ‘become a girl,’ while Sahu *mosolese* also denotes the “the wife of a genealogically younger member of a hereditary group.” Kashaya *cap<sup>h</sup>ya* also means ‘celibate.’ Oneida *yáh tha?tewa?alyáku* contains the negative particle *yáh* and the lexical roots *-a?al-* ‘net, lace,’ and *.ya?k-* “detach, sever, break, cut in two.” San Lucas Quiaviní Zapotec *digàì* is also a name for the ‘five centavo coin,’ Wintu *loymes* is derived from *loy* ‘put on front apron, become an adolescent girl,’ ‘virgin spirit, faery,’ while Chayahuita *nanon miáchin* contains *nanon* ‘girl’ and *miáchin*, which means ‘a bit, rather’ when following an adjective or adverb and ‘much’ otherwise. Great Andamanese *ôtlêkinga* also means ‘poor,’ while Hawaiian *pu’u-pa’a* is analyzable as ‘mound-firm’ and also denotes the ‘kidneys’ and, figuratively, ‘emotions, affections.’

#### 160. The Widow

Representation: 80%

Motivated: 32.1%

Thereof Analyzable: 26.9%

Thereof Colexifying: 5.2%

Thereof by Contiguity: 22.3%

Thereof by Similarity: 1.5%

Recurrent associated meanings: woman, husband, die/dead, orphan, unmarried, NEG, single, prostitute, lose, grieve/mourn, alone

Many of the motivated terms for this meaning are complex and of the lexical type, with one constituent being ‘woman,’ and the other indicating the differentia specifica. A common subtype is that making reference to the fact that a widow has ‘lost’ her husband, that is, that he has ‘died,’ as for instance Hawaiian *wahine kâne make* ‘woman husband die,’ Tsafiki *puyamin sono*, presumably containing *puyano* ‘to die’ and *sono* ‘woman,’ or Chickasaw *ihoo hattak imilli-* ‘woman man lose-NMLZ.’ Such terms are also featured in Efik, Katcha, Lesser Antillean Creole French, and Samoan. This particular denomination strategy need

not always be realized by terms containing a lexical element meaning 'woman:' next to derived terms in Carib and Manange, Kanuri has *kám-bà* 'person/man-not,' Itzaj *ix-kimen-icham* 'FEM.NOUN.CLASS-dead-husband' and *ix-ma'-icham* 'FEM.NOUN.CLASS-NEG-husband,' Kiliwa *?kuswa?=p-i?+hiw* 'husband=MP-DIST+fly,' Lake Miwok *míiw-helak* 'husband-lack,' and Cavineña *eahue-maju-que* 'husband-die-who.' Similarly, Baruya has *kwaimaaya* /kwalai-maaya-da' ~ kwalai-maaya-sa'/ 'man-without-she' and Rotokas *oira asava*, containing *asa* 'without.' The association with 'dying' is also found in Japanese (*mi-bō-jin* 'not.yet-die-person'), Cahuilla (*múk-vel* 'get.sick.or.weak/die-ABS.NMLZ'), and there are semianalyzable term where the identifiable constituents are verbs meaning 'to die' in Piro and Lenakel.

An association present also a few times in Indo-European (Buck 1949: 131) is that with the meaning 'alone' (as is the case in Old Norse and some of its daughters, Buck 1949: 131): Pawnee has *capaktihuks*, literally 'alone woman' (which is also quite literally used for a woman who lives alone), Bororo *aredu koadureudo* 'woman alone/widowed,' Burarra colexifies the relevant meanings, and there are semianalyzable terms where the identifiable constituent is 'alone' in Blackfoot and Yanomámi. In the latter language, the relevant term colexifies 'widow' with 'deprived of a loved one or goods,' 'with empty hands,' and 'orphan,' colexification with 'orphan,' as in one case in Indo-European according to Buck (1949: 131), is also encountered in Rendille, inter alia in Abzakh Adyghe, Chickasaw, and Wintu, where the relevant term *lolcit* also may refer to anyone in mourning, a 'survivor,' and someone possessed by the ghost of the deceased husband or wife, and hence 'cursed;' indeed, it is related to the root *lol* 'bereaved, ghost of dead spouse or divorced spouse; orphan(ed), phantom.' Similarly, Rotokas has *virakoiva*, with the gender/number marker *-va* 'feminine singular' for 'widow' and *virakoito* with *-to* 'masculine singular' for 'orphan,' a case of grammatical alternation. Toaripi has *ua lelesi* 'woman unmarried,' denoting "a woman without a husband (for any reason)" (which is also said of the relevant term in Swahili) and 'widow' and 'unmarried (woman)' are colexified in Efik, Buin, and Fijian. Khoekhoe has *loa-tara-s* 'grieve-woman-3SG.FEM,' and Biloxi *a'xti a"tcode" /a'xti' a"-tcode"/* 'woman mourn.' There are also other complex terms with one constituent being 'woman': Mbum has *wúŋ yìù* 'woman celibate' (this term also sometimes refers to a "prostitute or widow who accept young boys to sleep with," 'widow' and 'prostitute' are also colexified in Efik) and *wúŋ kúl* 'woman single' (compare colexification of 'single' and 'widow' in Welsh), Kwoma has *mibiya mima*, where *mima* is 'woman' and *mibiya* refers to the state of having been married before but not anymore, due to death of the spouse or divorce, Kyaka has *enda waiya* (*petenge*) 'woman chopped (living),' *enda mee peta-mo doko* 'woman empty sticking-ASSOC that' (this term has the variant *etembo peta-mo* 'single sticking-ASSOC'), Kosarek Yale *youwok kelabo* /youwi-ok kelabo/ 'sterile-just/alone women,' Chukchi *janraŋaw* /janra-ŋew/ 'separate-woman' (note that the root *\*weidh-* 'separate' is the origin of some Indo-European terms for 'widow,' Buck 1949: 131), Nivkh *tyrmu-umgu* 'lonely woman,' Cheyenne *otôxaa'é'e* perhaps literally 'uncovered woman,' San Mateo del Mar Huave *lemben omal najtaj*, seemingly containing *omal* 'point' alongside *najtaj* 'woman,' Santiago Mexquititlan Otomí *ra-nxu* 'one-woman,' Hani *miqcyuq* /almiq-cyuq/ 'woman-poor,' and White Hmong *poj-ntsuam* 'woman-destitute.one.' In addition, there are semianalyzable

terms with a constituent meaning ‘woman’ and/or ‘wife’ in Berik, Dadibi, Japanese, Highland Chontal, Cubeo, Miskito, and Great Andamanese.

Other associations include: Buli *pokong* is also the name of a particular tree and its fruit, Hausa *bazawara* ~ *zawara* also denotes “anything which has been used and is to be re-used” inter alia, while Noni *kpwεε ŋkfu* appears to contain *ŋkfu* ‘late.’ Buin *rarupere* also means ‘banana garden,’ Burarra colexifies ‘widow’ with “divorced or lone person,” Meyah *óna mesina* ‘widower’ is analyzable as ‘male string.bag,’ and Nunggubuyu *n<sup>a</sup>ayi* also denotes “any close kin of dead person” generally. Muna *kowaluno* ‘widow in mourning period’ is related to *walu* ‘to shroud’ and *kowalu* ‘shroud,’ since in Muna culture widows are wearing a white sarong over their heads in the 100 day mourning period they undergo. Ngaanyatjarra *wanakaarla* may be related to *wana* ‘digging stick’ (compare Dixon 2002: 99 for this association in Australian languages), Yir Yoront *thum-kuwn(l)* appears to be analyzable as ‘fire-grandchild,’ and *yoq-warry+pann* is analyzable as ‘tree-poor.fellow+younger.sibling.’ Abzakh Adyghe *px<sup>o</sup>əž* is analyzable as /*px<sup>o</sup>ə-žə*/ ‘girl-old,’ and Badaga *kunḍe* colexifies ‘widow’ with ‘woman’ generally (though this is archaic), and *munḍe* also means ‘bad woman, bad wife.’ Ineseño Chumash *’unitaxiš* contains *’unitax-* ‘to leave behind,’ while Wappo *kà· hanchóya*, containing *ká·* ‘person,’ also may refer to a “lone person, person without friends or family, poor person.” Central Yup’ik *aipaineq* is analyzable as /*aipaq-(ng)ite-neq*/ ‘partner/mate/spouse-lack-thing.that.results.from.’ Huambisa colexifies ‘widow’ with ‘sister-in-law/brother-in-law,’ while Fijian *dawai* also denotes “one who has none to care for him and her,” and Malagasy *mpitòndra tèna* is literally ‘one who carries oneself.’