Mothers or sisters? The encoding of morphological knowledge

Audring, J.

Mothers or sisters? The encoding of morphological knowledge

Jenny Audring
Leiden University

Abstract

How is grammatical knowledge encoded in mental representations? While traditional accounts view grammar as a system of rules, construction-based theories assume declarative schemas – lexical entries with variables – as the locus of grammatical knowledge. Such schemas are evidently needed to encode productive patterns. However, morphological knowledge also includes relations between existing words, in patterns that cannot necessarily be productively extended. This contribution argues that such patterns can be encoded in two ways: by a 'mother' schema dominating the listed instances, or by 'sister' links between the instances themselves. Sister links are the more parsimonious option, since they do not require a superordinate layer in the constructional network. However, mother schemas can encode properties that sister links cannot. This paper aims to work out how the division of labour between sister links and mother schemas may be organized.

Keywords: morphology, lexicon, paradigmatic relations, sister links, second-order schemas

1. Introduction

In construction-based theories, grammatical knowledge is usually seen as captured in lexical entries with one or more variables. For example, the morphological structure of the English words *boyish*, *foolish* and *childish* can be generalized as in (1).

(1) $[\text{N -}\textit{ish}]_{\text{A}}$ 'like N'

This entry, which says that English has adjectives consisting of a nominal base plus the suffix *-ish* and meaning 'like N', is commonly represented as a 'mother' node superordinate to its 'daughter' instantiations, as in (2).

Word Structure 12.3 (2019): 274–296
DOI: 10.3366/word.2019.0150

www.euppublishing.com/word

(2) [N-*ish*]$_A$ 'like N'

[*boy* -*ish*]$_A$ 'like a boy' [*fool* -*ish*]$_A$ 'like a fool' [*child*-*ish*]$_A$ 'like a child'

In this article I will ask whether this is the only way, or (necessarily) the best way, to represent linguistic generalizations. I will develop an alternative, the notion of *sister relations*, and discuss the possible division of labour between these and the traditional mother-daughter relations. The paper is theoretical in nature and employs a model called *Relational Morphology* (RM; Jackendoff & Audring 2016, 2019 and in press). RM is closely related to Construction Morphology (Booij 2010), differing from Booij's model in nuances rather than in substance. The main differences can be summarized as follows:

a. RM places a stronger emphasis on the declarative nature of constructional knowledge and the relations between existing lexical items.
b. RM uses a more detailed formalism that strictly separates morphosyntactic structure from phonological and orthographic structure, in line with the principles of the Parallel Architecture (Jackendoff 1997, 2002); see §2.2 for details.
c. RM does not privilege form-meaning pairings, but allows for meaningless constructions such as phonology-orthography mappings or purely morphosyntactic entities such as V or NP.

The paper starts with a brief outline of theoretical preliminaries. §3 sketches the difference between generalizations encoded in mother nodes and generalizations in the form of sister links. §4 applies the two types of generalization to various kinds of morphological relations and works out the virtues and limitations of each solution. §5 offers a conclusion and a brief outlook. Empirical support is drawn from West Germanic, specifically English, Dutch and German.

2. Theoretical preliminaries

Construction-based theories come in a variety of flavours, not all of them subscribing to the same commitments. One commitment that is essential for the issues at hand will be briefly reviewed here: the commitment to a *full entry* model of lexical items.

2.1 Full entry

Construction-based theories are principally non-derivational and assign a central role to the storage of linguistic structures in memory. Construction Morphology (Booij 2010, 2017, Masini & Audring 2019) is particularly explicit about the fact that such a theory requires a *full entry* type of architecture (Jackendoff 1975; Hilpert 2014: 66 uses the term "redundant representations"). This means that all lexical entries are fully specified, even when (parts of) their form and/or meaning recur in other lexical items. This stands in contrast to an *impoverished entry* type of architecture ("complete inheritance" in

Hilpert's terms), where entries only contain idiosyncratic information, plus pointers to other entries that encode more general properties. General information is considered to be higher up in the network; specific information is seen as lower down. Impoverished entries are incomplete, but storage-efficient, as every piece of information is listed only once. All information is encoded on the highest possible level. The mechanism that completes impoverished entries with the help of other, more general entries is usually called *inheritance*.[1]

Example (3) illustrates the difference between full entries and impoverished entries with the help of the English adjective *sheepish*. This word has the regular structure [N -*ish*]$_A$ and predictably means 'like a sheep'. In a network where both the base noun and the [N -*ish*]$_A$ schema have their own representations, an impoverished entry for *sheepish* would omit this information (the crossed-out parts in (3)), since it can be inherited from these representations. The entry for *sheepish* would then only list the unpredictable aspect of the word's semantics, namely the idiomatic meaning 'bashful'.[2]

(3) [N-*ish*]$_A$ 'like N' *sheep*$_N$ 'sheep'

[~~*sheep*$_N$ -*ish*~~]$_A$ '~~like a sheep,~~ bashful'

By contrast, a full entry representation of the same three lexical items would contain all information about *sheepish*, including those parts that are redundant with [N -*ish*]$_A$ and *sheep*$_N$.

Jackendoff & Audring (in press) discuss theoretical and empirical problems with impoverished entry theories and sketch the advantages of a full entry approach. For present purposes, full entry is important because it offers certain modelling options. Central among these is linkage by shared structure, as will be shown next.

2.2 Shared structure: sameness and equivalence

Lexical items in a full entry model can be connected by virtue of the structure they share, whereby "structure" is taken to mean any kind of linguistic property, be it phonological, morphological or semantic. This is clearer if we make the notation more precise. The adjective *boyish* and the corresponding affix schema can be unpacked as follows (using the notation of Jackendoff & Audring 2016, 2019 and in press).

(4) a. Semantics: $[\text{LIKE}_2\ (\text{BOY}_1)]_3$
Morphosyntax: $[_A\ \text{N}_1\ \text{-aff}_2]_3$
Phonology: /bɔɪ$_1$ ɪʃ$_2$/$_3$

[1] See Booij (2017) and Jackendoff & Audring (2016 and in press) for a reconceptualization of inheritance in a full entry model.

[2] One might argue that 'bashful' is really the only meaning of the word. However, the simile is part of the expressive power of the word and therefore cannot be ignored.

b.	Semantics:	$[\text{LIKE}_2\ (\text{X}_x)]_y$
	Morphosyntax:	$[_A\ \text{N}_x\ \text{-aff}_2]_y$
	Phonology:	$/\ldots_x\ \text{ɪʃ}_2/_y$

This notation follows the principles of the Parallel Architecture (Jackendoff 1997, 2002) by representing lexical items as a complex of meaning, morphosyntax and sound. Not every tier needs to be present, and other tiers, such as orthography, can be added. The structures on each tier are connected by *interface links*. The graphic representation represents these links by subscript coindices specifying what is connected with what. We use letters to link variables and numbers to link constants.[3] Thus, index 1 links the semantics, morphosyntax and phonology of the word *boy* in (4a), while 2 does the same for the affix and 3 for the entire complex word *boyish*.[4] The affix schema in (4b) has the same structure, with the exception that it contains a *variable* marked by the index x. This variable corresponds to the base noun *boy* in (4a) as well as to any other noun occurring with the suffix *-ish*. The schema as a whole is held together by the variable index y.

Recasting the mini network from (2) above in the more elaborate notation in (5) below, we see that the indices have a second use: in addition to connecting the tiers of a single item as interface links, they also function as *relational links* by connecting shared parts between two or more items. In this way, they specify what the lines in (2) express informally.

(5)

	$[\text{LIKE}_2\ (\text{X}_x)]_y$ $[_A\ \text{N}_x\ \text{-aff}_2]_y$ $/\ldots_x\ \text{ɪʃ}_2/_y$	CHILD_6 N_6 $/\text{tʃaɪld}/_6$
$[\text{LIKE}_2\ (\text{BOY}_1)]_3$ $[_A\ \text{N}_1\ \text{-aff}_2]_3$ $/\text{bɔɪ}_1\ \text{ɪʃ}_2/_3$	$[\text{LIKE}_2\ (\text{FOOL}_4)]_5$ $[_A\ \text{N}_4\ \text{-aff}_2]_5$ $/\text{fu:l}_4\ \text{ɪʃ}_2/_5$	$[\text{LIKE}_2\ (\text{CHILD}_6),\ \text{IMMATURE}]_7$ $[_A\ \text{N}_6\ \text{-aff}_2]_7$ $/\text{tʃaɪld}_6\ \text{ɪʃ}_2/_7$

To be more precise, (5) shows two types of shared-structure links. The words *boyish*, *foolish* and *childish* and the $[\text{N}$ *-ish*$]_A$ schema share the suffix *-ish*. Here, shared structure actually means *same* structure, indicated by the constant index 2. Another same-structure link connects the noun *child* and the base of the adjective *childish*, indicated by the constant coindex 6. Moreover, *boyish*, *foolish* and *childish* and the $[\text{N}$ *-ish*$]_A$ schema share the category A and the base category N. Here the shared structures are not the same but *equivalent*, since the words contain specific nouns whereas the schema contains a nominal variable. Translated into the formalism, this means that the constant indices 1, 4 and 6 in the words map onto the variable index x in the schema, while the constant indices 3, 5 and 7 map onto the variable index y.

[3] Interface links work the same for variables and constants, so this notational difference is of no consequence here. However, it matters for *relational links* and the difference between same structures and equivalent structures introduced in the following paragraph.

[4] The placement of index 2 on the semantic tier is a simplification. See Jackendoff & Audring (in press) for a discussion of the complexities.

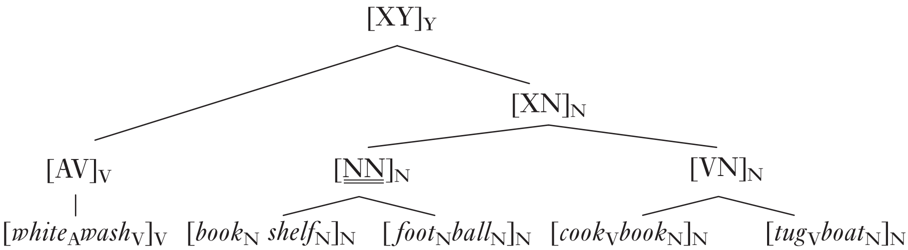
Importantly, linkage by shared structure requires redundancy; after all, shared structure means structure represented (at least) twice. Therefore this option is not available in impoverished entry models where all redundant structure is eliminated. In a full entry model, linkage by shared structure is a powerful tool with a high potential of psychological plausibility; after all, associative connections between similar items are widely accepted in memory research. Construction-based theories that embrace a full entry architecture are perfectly compatible with this type of memory model.

Linkage by shared structure is not the only advantage of a full entry approach. Another is the absence of a need to prioritize higher nodes in the taxonomy, as will be briefly discussed next.

2.3 Local generalizations

Freedom from the commitment to "save space" by encoding properties only once at the highest possible level has the consequence that higher nodes in the hierarchy are not necessarily a priority. Instead, generalizations can be stated on every level of the hierarchy, including low-level generalizations capturing "local" insights about smaller groups of words. In a model where every node is fully specified, higher nodes are if anything less informative, as every step up in generality removes specific information (Goldberg's *surface generalization hypothesis* (2002: 329) can be interpreted as making the same point; see also Hilpert 2019 (in this issue)). For example, consider the (partial) taxonomy of compounds in (6), from the specific words at the bottom to the maximally general schema at the top (for the sake of readability, (6) is given in shorthand notation; see Langacker 2008: 237ff. for a similar schema and argumentation).

(6)



The first level of generalization encodes the fact that English has, among others, verbal AV compounds and nominal NN and VN compounds (note that this is presented as the lowest level of generalization about compounds, but it makes use of categories like A, V and N which themselves are high-level abstractions). The double-underlined schema $[\underline{\underline{NN}}]_N$ is productive: new NN compounds can be constructed freely and with ease. The $[XN]_N$ schema on the second level says that the language has nominal compounds, and the general schema $[XY]_Y$ states that the language has compounds of some kind. Every entry, from the lowest to the highest level, furthermore encodes the fact that the English compounds in (6) are formally right-headed; this is shown in the correspondence between the category of the right-hand part and the category of the entire complex word (for NN compounds, right-headedness can be established with the help of other characteristics, such as the shape and position of the plural marker).

Two consequences of this setup are noteworthy. First, it offers the freedom to make abstract statements at exactly the degree of generality that is supported by the data. For example, we can state the productivity of NN compounds on the level where it applies. Any higher-level node would contain conflicting information about productivity. For example, $[XN]_N$ is fully productive if X is a noun, but less productive if X is a verb (and matters get complicated if X is an adjective, as AN compounds are notoriously hard to distinguish from AN phrases in English, and new forms may arise by stress shift from phrases rather than by productive compounding (Bauer, Lieber & Plag 2013: 451–452)). In fact, productivity might need to be stated yet lower in the hierarchy, as unproductive patterns can contain pockets of productivity and vice versa (see e.g. Hüning 2009 and Cappelle 2014). For example, the productivity of VN compounds is restricted, and there are preference patterns both for the semantic type of the noun and for the form of the verb (Gast 2008). Certain verbs enter such compounds in the plain stem form, as in *payday, payroll, paycheck* and *payphone*, others as a gerund, as in *reading room, reading glasses, reading week* and *reading lamp*. Exocentric VN compounds like *spoilsport* or *pickpocket* have quirks of their own. The network allows for such patterns to be stated locally, superseding more general information encoded on higher levels.

The second consequence of the arrangement in (6) is that the higher levels are increasingly redundant with respect to the lower. Indeed, one may wonder what the usefulness of the higher schemas might be. Even if it were true that all compounds in the language were consistently right-headed (which is not the case for English), this information would be cumulatively expressed by the lower-order schemas. It is not so clear what process would need to access this information at the highest level.[5] The main motivation for stating it might be the fear of losing a maximal generalization, but the necessity of a maximal generalization can be questioned. This stands in stark contrast to an impoverished entry model of the same compound network, where the highest node alone would represent the headedness parameter (better known in morphology as the Righthand Head Rule; Williams 1981) and therefore be of paramount importance.

We can conclude that in construction-based theories, higher-level generalizations do not enjoy an automatic privilege. They might either fail to arise in the first place, or sink into oblivion if they are not regularly called up in the course of language processing.[6] The circumstances under which a higher-level schema might be expendable are the central concern of this paper.

[5] From a diachronic point of view, the tendency of languages to develop a general preference with regard to headedness could be the cumulative result of countless instances of low-level choices to conform to existing words and patterns rather than arising from the pressures of a very high-level schema saying where heads should be located.

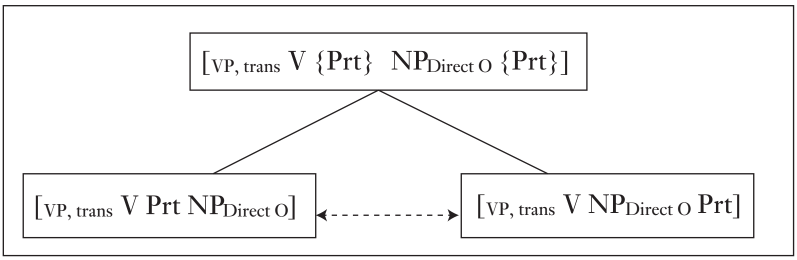
[6] This perspective ties in with the understanding, spreading within linguistics, that many grammatical patterns fail to apply at a truly general level and are tied to individual surface forms (see Diessel 2016 for a summary and references). Consequently, every generalization should be expressed at the level where it holds, and the top levels of the generalization hierarchy may end up sparsely populated.

3. Mothers or sisters?

In the representations (2), (3), (5) and (6) above, all generalizations are expressed as mother–daughter links: a mother schema dominating two or more daughter words. This is a widely accepted setup in Construction Grammar (see e.g. Hoffmann 2017) and is also adopted in Construction Morphology (e.g. Booij 2010: 26). However, recent contributions have suggested the importance of *horizontal*, *lateral* or *paradigmatic* links, i.e. of *sister* relations; notable examples are Cappelle (2006), Booij (2010), Van de Velde (2014), Diessel (2015), Booij & Masini (2015), again Hoffmann (2017), Norde & Morris (2018) and Zehentner & Traugott (forthcoming). They are also of primary importance in exemplar-based approaches, e.g. Bybee (2010, 2013), which group entries by similarity and assign schemas with variables a secondary role.

Sister relations hold between items on the same level of the hierarchical network, i.e. between words and phrases of equal specificity and/or complexity.[7] They are often employed for alternations, i.e. for constructions that differ in certain pertinent semantic and/or formal aspects but are highly similar otherwise. If the contrast between the alternants is primarily formal and goes with a common meaning, horizontal relations are often understood as an addition to existing inheritance relations. For illustration, see Cappelle's treatment of particle placement, schematized in (7) (from Cappelle 2006: 18). The two syntactic options, one with the particle in between the verb and the direct object (*she turned off the TV*), and one with the particle at the end (*she turned the TV off*), are presented as *allostructions*, i.e. as variants of a superordinate construction much in the same way that allophones are variants of the same phoneme and allomorphs are variants of the same morpheme. In addition, the two allostructions have a lateral or sister relation, represented by a dashed line.

(7)



[7] The issue of complexity is relevant here because some sources (e.g. Booij 2010: 26) represent simplex words as 'parents' of complex words, the way *sheep* in (3) is represented as superordinate to *sheepish*. This representation fits naturally in a taxonomic hierarchy, in which complex words inherit the properties of their bases, since inheritance proceeds metaphorically 'downward'. The opposite arrangement is found in compositional hierarchies, i.e. in morphological and syntactic tree structures, where simplex elements constitute the lowest level.

Note that terms like 'up' and 'down', like 'mother' and 'daughter', should never be regarded as anything but metaphors when speaking of a mental network, but the issues are pertinent in a more than metaphorical way. If inheritance is theoretically restricted to information transfer from superordinate to subordinate nodes, then it matters which node is superordinate (the mother) and which is subordinate (the daughter).

In light of the considerations in §2, two questions arise. First, what is the status of the mother-daughter links and the sister link in (7)? The fact that the sister link is represented by a dashed line suggests a secondary connection. Indeed, Cappelle states that the two allostructions are linked "via a common 'supercategory'" (2006: 19); hence the relation is primarily by route of the mother. This view can be questioned. And second, we can inquire about the functionality of the mother. In the following, we will briefly run through the arguments, before discussing them in more detail for morphological patterns.

Following the argumentation in §2.2 above, the linking lines in (7) can be understood as relations based on shared structure. We could then recast (7) as (7'), with coindices in addition to lines (the specification "Direct O" is omitted in the interest of readability). All three items are schemas with variables, so the indices are given in letters (cf. (4) and (5) above for the use of letter and number indices).

(7') $[_{\text{VP, trans}}\ \text{V}_a\ \{\text{Prt}\}_b\ \text{NP}_c\ \{\text{Prt}\}_b]$

$[_{\text{VP, trans}}\ \text{V}_a\ \text{Prt}_b\ \text{NP}_c]$ ——— $[_{\text{VP, trans}}\ \text{V}_a\ \text{NP}_c\ \text{Prt}_b]$

The crucial consideration is that the verb, the noun phrase and the particle in the three nodes are the *same*: the schemas express two placement options for the same particles in the same syntactic configuration. Therefore, V, NP and Prt receive the same index. (7') then abounds in redundant information: each index appears at least three times.

From the traditional perspective favouring higher-level nodes, this might be seen as a redundancy of the daughters with respect to the mother. In view of §2.3, however, it makes equal sense to say that the mother is redundant with respect to the daughters. In fact, it makes better sense, since the mother is a generalization over the daughters, so the daughters are epistemologically primary.[8] Now an interesting option arises. Since the coindexation specifies that the constituents of the two allostructions are the same, the generalization encoded in the mother – namely that the same particle can appear in two places – is already expressed in the sister links between the daughters. This opens the door to the representation in (7'').

(7'') $[_{\text{VP, trans}}\ \text{V}_a\ \text{Prt}_b\ \text{NP}_c]$ ——— $[_{\text{VP, trans}}\ \text{V}_a\ \text{NP}_c\ \text{Prt}_b]$

[8] Here the metaphor breaks down, as the sisters give rise to the mother rather than vice versa. Also, there is a scenario in which the mother is discovered first, namely via individual schemas with variable particle placement, e.g. [*look* {*up*} *the word* {*up*}] and [*shut* {*down*} *the computer* {*down*}]. However, this analysis would not give us the lower-level allostructions in (7)–(7''), as these do not generalize over pairs like *look up the word* and *look the word up*, but over sets like *look up the word* and *shut down the computer*. In what order learners build up their knowledge is, of course, an empirical question, the answer to which may differ from speaker to speaker.

The lateral connection between the two allostructions now expresses the generalization directly by saying that the same particle (index b) can appear in two places relative to the same verb (index a) and the same NP (index c), namely between V and NP or after NP. No information is lost compared to (7'). Preference conditions for one schema or the other can still be stated for each of the sisters, as in the original proposal. On the general assumption that any stipulated entry should be useful, this means the mother is expendable.[9]

The upshot is that generalizations can be encoded in two different ways: either by means of a dedicated mother node or by enriching the relevant lexical items with sister links. This raises the question of what the division of labour might be between mother schemas and sister links. Under what circumstances is a mother advantageous or needed? When might sister links be sufficient?

4. Generalizations in morphology

Before tackling this question for morphology, it is useful to ask what kinds of generalizations morphology, the grammar of words, should be able to express. Simply speaking, generalizations can be drawn whenever two or more words are in some way *the same*, whereby sameness in morphology canonically involves all structural tiers. Thus, *fool* and *foolish* provide a better basis for a generalization than *van* and *vanish*, because the former share parts of their phonology, morphosyntax (of the base) and semantics, while the latter only share phonology. Also, the pattern of sameness in *fool* and *foolish* reappears in other word pairs, e.g. *child – childish* or *baby – babyish*, whereas *van – vanish* show a one-off relation. Systematic sameness, then, can be assumed to be more noteworthy than accidental sameness, which is why generalizations make better sense for structure shared across multiple tiers and multiple lexical items.

In the following sections, we will consider two basic types of systematic sameness, subsumed under the labels *same affix* and *same stem*. *Same affix* is the relation between, say, *childish* and *foolish*; *same stem* is the relation between word pairs like *construction* and *constructive*. As we will see, these are not entirely equivalent, since they involve different types of mother – a schematic mother for same-affix relations and a lexical mother for same-stem relations. Yet, they both contribute to elucidating the issues at hand.

4.1 Same affix

A common way to express same-affix generalizations is illustrated in §1, example (2). Elaborating (2) in the more precise notation yields (8), partly repeating (5) in §2.2.

[9] See also Booij (2010: 135) for an account of preverb/particle incorporation in Dutch, where separable complex verbs are modelled as two linked schemas without a superordinate structure.

(8)

$[\text{LIKE}_2\ (\text{X}_x)]_y$
$[_A\ \text{N}_x\ \text{-aff}_2]_y$
$/\ldots_x\ \text{ɪʃ}_2/_y$

$[\text{LIKE}_2\ (\text{BOY}_1)]_3$	$[\text{LIKE}_2\ (\text{FOOL}_4)]_5$	$[\text{LIKE}_2\ (\text{CHILD}_6), \text{IMMATURE}]_7$
$[_A\ \text{N}_1\ \text{-aff}_2]_3$	$[_A\ \text{N}_4\ \text{-aff}_2]_5$	$[_A\ \text{N}_6\ \text{-aff}_2]_7$
$/\text{bɔɪ}_1\ \text{ɪʃ}_2/_3$	$/\text{fuːl}_4\ \text{ɪʃ}_2/_5$	$/\text{tʃaɪld}_6\ \text{ɪʃ}_2/_7$

The coindices spell out the relational links between the three words and the [N *-ish*]$_A$ schema generalizing over their morphological structure. The question now arises whether such generalizations could be expressed by sister links alone, as suggested for the particle alternation in (7''). To this aim, it is necessary to inquire whether the mother has any additional functionality.

For the *-ish* adjectives, at least three potential advantages of a mother can be identified. A first difference between these adjectives and the particle alternation in (7) is the level of generality and, concomitantly, the number of the daughters. Since the particle alternation is represented on a higher schematic level, it only has two variants: the two allostructions. The [N *-ish*]$_A$ schema, by contrast, is a lower-level schema instantiated by 84 lemmas according to CELEX (celex.mpi.nl, searched July 27, 2018), excluding language or population names like *Finnish*, *Polish* or *British*, which would swell the group substantially. Assuming that a pattern is more conspicuous – and hence more likely to be elevated to a schema – the more daughters it has, a generalization such as [N *-ish*]$_A$ in (8) is more likely to have its own entry than a generalization such as $[_{\text{VP, trans}}\ \text{V}_a\ \{\text{Prt}\}_b\ \text{NP}_c\ \{\text{Prt}\}_b]$ in (7'). This consideration applies more widely: since taxonomic hierarchies are broader at the bottom, with more daughters on lower than on higher levels, lower-level mothers in general should be more likely than higher-level mothers. This ties in with the considerations about local generalizations discussed in §2.3.

The number of daughters, i.e. the type frequency, also matters for schemas on the same level of generality. Encoding the family ties within a large family like [N *-ish*]$_A$ with only sister links would mean linking every *-ish* word to every other *-ish* word, a huge proliferation of links. A mother schema like [N *-ish*]$_A$ can serve as a relational hub, allowing words to link to the mother instead of to all the sisters. It follows that a mother node is more functional, the larger the family of words it connects (logically speaking, the advantage should begin at four family members, as connecting every member to the mother would require four links instead of the six needed to connect every sister with every other sister; of course we have no way of knowing if the mind calculates cost in such a way). Thus, a very minor affix family like that of the German suffix *-icht*, which has only five low to very low frequency instances (*Dickicht* 'thicket', *Röhricht* 'reed bed', *Tannicht* 'fir forest' (archaic), *Kehricht* 'sweepings' and *Spülicht* 'dishwater'), reaps less benefit from a mother schema than a larger and more salient affix family like the English *-ish* adjectives – in fact, most speakers may not be aware of it at all.

A second obvious benefit of a mother schema is that it can encode productivity. As argued in Jackendoff & Audring (2016), productivity can be construed as the degree of openness of the schema's variable: the more open a variable is, the more easily it accepts new lexical material, which makes the pattern synchronically productive. In the case of the $[\text{N -}ish]_{\text{A}}$ schema we find novel formations such as *Trumpish*; hence, the N represents an open variable.[10] If productivity is encoded in the variable of the mother schema, then every systematically productive pattern needs a schema with a variable.[11] This is an issue that did not arise with the higher-level particle schema, as its daughters in (7) to (7'') are themselves schemas with their own degree of productivity, as indeed Cappelle (2006) shows.

For a third argument in favour of a mother schema in (8), consider the bases of the three *-ish* words. They all serve the same purpose in the complex structure, namely as a stem in morphosyntax and as an argument in the semantics. Simply put, *boy* has the same function in *boyish* as *child* has in *childish*. This shared function cannot be encoded by means of sister coindexation because the nouns themselves are different for each sister. Recalling §2.2, the relation between the bases is not one of sameness, but of equivalence. Example (8) captures this relation in the variable marked x, which maps unto the individual stems indexed as 1, 4 and 6. Thus, equivalence is a relation that cannot be expressed in the daughters themselves, but requires a mother to state.

This situation did not arise in (7) because the particle alternation was viewed in schematic form, so all schemas involve the same Prt, V and NP. When fully specified instantiations are added, as in (9), the two mother nodes show the same additional generalization over their daughters: 8 and 11 map onto a, 9 and 12 map onto b, 10 and 13 map onto c.

(9) $[_{\text{VP, trans}}\ \text{V}_a\ \text{Prt}_b\ \text{NP}_c]$ $[_{\text{VP, trans}}\ \text{V}_a\ \text{NP}_c\ \text{Prt}_b]$

$[\textit{turn}_8\ \textit{off}_9\ [\textit{the TV}]_{10}]$ $[\textit{look}_{11}\ \textit{up}_{12}\ [\textit{the word}]_{13}]$ $[\textit{turn}_8\ [\textit{the TV}]_{10}\ \textit{off}_9]$ $[\textit{look}_{11}\ [\textit{the word}]_{13}\ \textit{up}_{12}]$

In short, mother schemas can encode structures that differ in substance but converge in other ways, e.g. in functionality.

Let us briefly consider a further argument in favour of a mother schema. If, as stated in §4, generalizations make better sense for structure shared not only across multiple items but also across multiple tiers, then a schema with semantic, morphosyntactic and phonological specifications would make a better mother than a schema with, say, only phonology. Hence, the schema for the English suffix *-ish* is more informative than a putative schema for the German suffix *-icht*, because the *-ish* adjectives in question all have the same base category (N), while the German words are deadjectival (*Dickicht*),

[10] The Oxford English Dictionary says: "In recent colloquial and journalistic use, *-ish* has become the favourite ending for forming adjectives for the nonce (esp. of a slighting or depreciatory nature) on proper names of persons, places, or things, and even on phrases" (www.oed.com, accessed 2019-08-11).

[11] The formulation "systematically productive" excludes cases of paradigmatic word-formation (discussed e.g. in van Marle 1985, Meesters 2004 and Booij 2019b for Dutch) where a new word is formed analogically to an existing one. Such cases make use of sister links, not a mother schema.

denominal (*Röhricht* and *Tannicht*) or deverbal (*Kehricht* and *Spülicht*). On the semantic side, the *-ish* family is united by the fact that most [N *-ish*] adjectives have a pejorative connotation and are built around nouns denoting humans or other animates – out of the 84 relevant lemmas in CELEX, at least 73 have an animate base.[12] The five German nouns, by contrast, fall into two unrelated semantic groups: *Dickicht*, *Röhricht* and *Tannicht* denote clusters of vegetation and *Kehricht* and *Spülicht* are mass nouns used in a household context.

However, the argument that more shared structure provides a better base for a generalization is not necessarily an argument for a dedicated mother schema. Instead it could be interpreted as increasing the likelihood that two (or more) words are recognized as sisters in the first place. Hence, more conspicuous similarity boosts the chances of a mother schema, but does not necessarily predict the rise of the schema itself. If we do interpret it in favour of a mother, it should be noted that there is a trade-off between family size and specificity of the mother schema: the more specific the schema, the fewer instances it has. Hence, it might make sense to expect a balance in terms of mental computation, preferring schemas that are sufficiently specific to be informative, yet inclusive enough to generalize over a substantial number of instances.

Summarizing the discussion so far, we have seen at least three situations in which a mother schema is likely or required:

– for patterns with a high type frequency, and for lower-level patterns in general;
– for productive patterns;
– for generalizations that cannot be encoded in the daughters, e.g. when structures are equivalent but not the same.

The third situation has other manifestations, as will be shown below.

4.2 Same stem

Let us next look at same-stem relations such as in *construction* and *constructive*, where the issues are more complex. Same-stem relations can be viewed on two levels: the level of the word and the level of the schema. We will start with words (§4.2.1) and then scale up to schemas (§4.2.2). Note that the term "stem" will be used rather loosely, including not only inflectional stems but also derivational bases.

4.2.1 Word-formation

In the simplest case, two or more words share a stem that corresponds to a free word. This word can then act as a mother node to connect the derivatives, the way *sheep* and

[12] The remaining adjectives with inanimate noun bases are *bookish, faddish, feverish, hellish, huffish, liverish, modish, nightmarish, novelettish* and *stylish*, plus *pettish* with a now obscure base. *Lumpish* and *cloddish* are ambiguous.

child were represented as mother nodes to *sheepish* and *childish* in (3) and (5) above. In example (10), *child* is the mother of the two derived words *childish* and *childless* – again, the shorthand notation suffices for now. Note that this is a special type of mother, as it is not schematic; we might call it a *lexical mother*.

(10) *child*$_{\mathrm{N}}$

[*child*$_{\mathrm{N}}$ *-ish*]$_{\mathrm{A}}$ [*child*$_{\mathrm{N}}$ *-less*]$_{\mathrm{A}}$

Again, the question arises if (10) could be expressed as a sister relation rather than a mother–daughter relation. Interestingly, the stakes are different here: the mother is an existing word and therefore part of the lexical network anyway; hence, the mother–daughter relation can be employed without additional justification for the extra node. The question then flips: the mother comes for free, but are there reasons to assume a sister relation as well?

Interestingly, the notation gives us the sister configuration automatically. If (10) is repeated in the more elaborate notation (see (11); the coindices correspond to (5) and (8) above), coindex 6 does not only relate the mother node *child* to its daughters, but also indicates a connection between the daughters.

(11) CHILD$_6$
N$_6$
/tʃaɪld/$_6$

[LIKE$_2$ (CHILD$_6$), IMMATURE]$_7$
[$_{\mathrm{A}}$ N$_6$ -aff$_2$]$_7$
/tʃaɪld$_6$ ɪʃ$_2$/$_7$

[WITHOUT$_{14}$ (CHILD$_6$)]$_{15}$
[$_{\mathrm{A}}$ N$_6$ -aff$_{14}$]$_{15}$
/tʃaɪld$_6$ ləs$_{14}$/$_{15}$

While this is an artifact of the notation, it does reflect the fact that the sister configuration should be easy to detect. Once the sameness between the mother and each daughter is established, the sameness between the daughters should also be evident. Encoding it in the network involves the small step of creating a link rather than the big step of creating a new schema. Hence, sister links can be considered a likely addition in a configuration like (11). On the other hand, as argued in §4.1, it may be that not all potential sister links are actually instantiated when the family gets large, e.g. when an affix is highly productive. In such cases, not every sister may end up linked to every other sister (see Jackendoff & Audring in press for a fuller discussion).

The situation is different for stems that do not have an independent existence. This is especially relevant for inflection, but it can also be observed in word-formation. Consider pairs such as *ambition – ambitious*, *splendid – splendor* or *hilarious – hilarity*. These words are clearly sisters, since they share a substantial amount of phonology and semantics, but their bases do not occur as free words. They carry no independent meaning, their category cannot be ascertained and even their phonology is not always unambiguous (is it *hilar-ity*, parallel to *timid-ity*, or *hilari-ty*, parallel to *royal-ty*?). Hence positing bound roots such as *ambi(t)-*, *splend-* and *hilar(i)-* as lexical entries is questionable: the only function they have is to represent the base of complex words that themselves have to be listed. This information, however, can also be encoded by coindexing the words themselves. Even in cases with more than two siblings, as in the family of *hor-* words (*horror*, *horrid*, *horrible*, *horrify*, *horrific*, *horrendous* and *abhor*), positing *hor-* as an independent lexical entry seems expendable.

Summarizing, there are two scenarios for same-stem relations. When the stem occurs as a free word, this word can serve as a lexical mother. When the stem does not occur in isolation, a sister configuration is more parsimonious.

The situation gets more interesting and complex if the same-stem relation between two words is non-unique. While *childish* and *childless* is the only word pair where *-ish* and *-less* occur with the same noun, a comparison of the suffixes *-less* and *-ful* yields 39 words with a shared base.[13] Even for bound stems, non-unique pairs can be found, e.g. *caution – cautious* and *ambition – ambitious*, or *candid – candor* and *splendid – splendor*. In patterns of this kind, the same-stem relation can extend from the word level to the schema level. Such cases will be discussed next.

4.2.2 Systematic same-stem relations: second-order schemas

Same-stem relations can be systematic to some degree, spanning smaller or larger groups of words. An example for a larger group is the relation between [N -*ful*]$_{\mathrm{A}}$ and [N -*less*]$_{\mathrm{A}}$. As mentioned in §4.2.1, the two patterns are linked by a fair degree of overlap in the list of attested bases. (12) sketches the relation.

(12)
$$\begin{tikzcd}[\mathrm{N}\ \textit{-ful}]_{\mathrm{A}} \arrow[r, dash] \arrow[d, dash] & [\mathrm{N}\ \textit{-less}]_{\mathrm{A}} \arrow[d, dash] \\ [\textit{help}_{\mathrm{N}}\ \textit{-ful}]_{\mathrm{A}} \arrow[r, dash] & [\textit{help}_{\mathrm{N}}\ \textit{-less}]_{\mathrm{A}}\end{tikzcd}$$

The same is true for [ø -*id*]$_{\mathrm{A}}$ and [ø -*or*]$_{\mathrm{N}}$ where the group is smaller and the base is a non-lexical root. Examples are given in (13) (from Jackendoff & Audring in press; ø stands for the root).

[13] Interestingly, not every N-*less* word has an N-*ful* sister. CELEX contains nearly twice as many words of the pattern [N -*less*]$_{\mathrm{A}}$ (N=200) as of the pattern [N -*ful*]$_{\mathrm{A}}$ (N=106). The numbers contain a considerable amount of noise, as the base category is often ambiguous and CELEX analyses are sometimes questionable (words like *watchful*, for example, appear as [*watch*$_{\mathrm{V}}$]$_{\mathrm{N}}$*ful*$_{\mathrm{A}}$, with the base analysed as a nominalized verb). However, the noise in the two lists is of the same kind, which should make the numbers comparable.

(13) $[ø\ \text{-}id]_A$ —— $[ø\ \text{-}or]_N$

candid	*candor*
fervid	*fervor*
horrid	*horror*
languid	*languor*
pallid	*pallor*
splendid	*splendor*
squalid	*squalor*
torpid	*torpor*
(possibly) *stupid*	*stupor*
(possibly) *liquid*	*liquor*[14]

The upper levels of (12) and (13) show a configuration that Booij & Masini (2015) refer to as a *second-order schema* (the term is adopted from Nesset 2008; see also Kapatsinski 2013; Plag 2003 refers to it as a *cross-formation*). Second-order schemas encode systematic relations between schemas (see also Booij 2019a (in this issue)).

For present purposes, an important question to ask is whether second-order schemas involve or constitute a mother schema superordinate to the two (or more) schemas participating in the relation. The representations in (12) and (13) show that this is not the case. The relation between the two schemas lies in the fact that they share the same variable for the base. Example (14), which is a fuller representation of (12), illustrates that this relation can be expressed by coindexing the base variables (index a).

(14)

$[\text{HAVING MUCH}_{16}\ (X_a)]_d$	$[\text{WITHOUT}_{14}\ (X_a)]_e$
$[_A\ N_a\ \text{-aff}_{16}]_d$	$[_A\ N_a\ \text{-aff}_{14}]_e$
$/\ldots_a\ \text{fəl}_{16}/_d$	$/\ldots_a\ \text{ləs}_{14}/_e$

Adding a mother to (14) would yield a schema with nothing but the category of the base (N) and the category of the complex word (A). Other schema pairs, e.g. $[ø\ \text{-}id]_A$ and $[ø\ \text{-}or]_N$ in (13), do not yield words of the same category, so a putative mother would be even less informative. In addition, the mother schema might say that the complex words contain an affix. However, this generalization only covers these two particular affixes, which would need to be listed as daughters. Under such circumstances, a mother schema contains no information that is not also contained in the daughters. Even the fact that

[14] In addition, there are cases that are diachronically related, but syncronically divergent, such as *rigid – rigor*, *humid – humor*, *rancid – rancor* and *valid – valor*. Spelling with *-or* is chosen here for the sake of uniformity; some (but not all) of the nouns have a variant in *-our* (*splendour* vs. **squalour*).

both nouns in (14) function as bases, which was identified as a motivation for a mother in §4.1, is captured by the sister index a. As a consequence, second-order schemas can be understood as a pair or group of *sister schemas*. They represent an extra degree of abstraction compared to the participating schemas in isolation, but they do so without an additional superordinate node.

Note that the situation can change if the schemas involved are united by additional properties. For example, the question arises how the antonymic relation of $[\mathrm{N}\ \textit{-ful}]_{\mathrm{A}}$ and $[\mathrm{N}\ \textit{-less}]_{\mathrm{A}}$ might be represented. One way (out of many) is to construe the opposition as endpoints of a scale ('degree of having N'). This scale would require a mother schema to encode.[15] What we see here is the same situation as in §4.1 above: mother schemas might be needed to capture a generalization that cannot be stated in the daughters. As above, such situations occur when two or more sister schemas share properties beyond actual structural overlap. This point will recur in the next section which discusses inflectional paradigms.

It is worth pointing out that a sister relation between schemas does not necessarily imply full productivity in both directions. English has *clueless* but not **clueful*, *gleeful* but not **gleeless*, *lucid* but not **lucor*, and *error* but not **errid*. On the other hand, there are sister schemas where having one sister implies a fair chance of also having – or being able to productively form – the other. The most famous case is probably the schema pair $[\mathrm{X}\ \textit{-ism}]_{\mathrm{N}}$ and $[\mathrm{X}\ \textit{-ist}]_{\mathrm{N}}$, as in *Marxism – Marxist* or *pacifism – pacifist*, discussed in Booij (2010: 31–36) (although cases of blocking occur: an adherent to *stoicism* is a *stoic*, not a **stoicist*).[16] Full mutual productivity between sister schemas is expected in inflection (see §4.2.3 below).

For our purposes, second-order schemas are particularly interesting because they combine the power of mother schemas and sister links. Each of the participating schemas is a mother to a family of individual words, but the relation between them is expressed by means of sister links. Moreover, the system of linkage is the same for sister words and for sister schemas, as shown in (12) (coindexation would connect *help* on the word level and the variable N on the schema level), which yields a homogeneous modelling solution for all levels in the constructional taxonomy.

Second-order schemas, i.e. sister relations between schemas, have also been proposed as a modelling solution for inflection (Booij 2010: 256, van der Spuy 2017, Masini & Audring 2019: 384). Yet, the theory is still in its infancy; especially paradigms have not been addressed in any depth or detail (see also Diewald to appear, who stresses the

[15] Thanks to Steffen Höder for pointing this out.

[16] Booij also points out that sister words can show a closer semantic connection to each other than to their base: "A Marxist is an adherent of Marxism and not necessarily a follower of Marx, since Marxism as a doctrine encompasses more than the ideas of Marx (in fact, Marx himself declared that he was not a Marxist). Similarly, a socialist is not necessarily a social person but an adherent of the ideology of socialism" (2010: 33). This is a compelling argument for a sister relation.

urgency of including paradigms in Construction Grammar, though her approach differs to some extent from the one developed here). The final section addresses a few of the issues.

4.2.3 Inflection

Same-stem relations hold derivational families together, but they play an even more central role in inflection. A crude representation of the singular indicative active paradigm of the German verb *machen* 'to make' is given in (15); using the same split between word level and schema level as in (12)–(14) above (the simplified notation omits mood and voice for readability's sake and neglects many complexities, such as syncretism). Members of a paradigm are represented as sisters, both on the lower level of the individual word forms and on the higher level of the schemas. Again, the schemas show a second-order relation, this time between three sister schemas, and more would be added for the full paradigm.

(15)
$$\begin{array}{ccccc}
[\mathrm{V}\ \text{-}e_{1\mathrm{SG.PRS}}] & \text{—} & [\mathrm{V}\ \text{-}st_{2\mathrm{SG.PRS}}] & \text{—} & [\mathrm{V}\ \text{-}t_{3\mathrm{SG.PRS}}] \\
| & & | & & | \\
[mach_{\mathrm{V}}\ \text{-}e_{1\mathrm{SG.PRS}}] & \text{—} & [mach_{\mathrm{V}}\ \text{-}st_{2\mathrm{SG.PRS}}] & \text{—} & [mach_{\mathrm{V}}\ \text{-}t_{3\mathrm{SG.PRS}}]
\end{array}$$

Is such a sister configuration sufficient to express paradigmatic relations in inflection, or is a mother schema needed? While the complexities are manifold, a few considerations can be offered.

First, let us consider what expressive power can be achieved with coindexation. As we saw in (14) above, the primary means to express a paradigmatic link between two or more sister schemas is the coindexation of the stems. For productive schemas, coindexing the stem variables encodes the fact that whatever fills the variable in one schema can also fill the variable in the other. In regular inflection, full productivity is generally the norm; hence, coindexing the V in each of the three schemas in (15) would amount to a valid statement, namely that verbs that end in *-e* in the first person singular end in *-st* in the second person singular and in *-t* in the third, and vice versa.

Different indices can be used for different inflectional classes. As an example, (16) represents two cells of the German nominal paradigm, the genitive singular cell and the nominative plural cell. Both suffixes have allomorphs, constituting a number of inflectional classes. (16) shows two of these classes, represented by the noun *Mensch* 'human, person' and the noun *Bild* 'picture, image'. Class membership is captured by coindices: index a connects the inflected forms of the $\text{-}en_{\mathrm{Gen.Sg}}/\text{-}en_{\mathrm{Nom.Pl}}$ class, index b the forms of the $\text{-}(e)s_{\mathrm{Gen.Sg}}/\text{-}er_{\mathrm{Nom.Pl}}$ class. The various forms within the same class are sisters, both on word and on schema level (only schema-level sisterhood is marked in (16)).

(16) a. $[\mathrm{N_a}\ \textit{-en}_{\mathrm{Gen.Sg}}]$ — $[\mathrm{N_a}\ \textit{-en}_{\mathrm{Nom.Pl}}]$

$[\textit{Mensch}_{\mathrm{N}}\ \textit{-en}_{\mathrm{Gen.Sg}}]$ — $[\textit{Mensch}_{\mathrm{N}}\ \textit{-en}_{\mathrm{Nom.Pl}}]$

b. $[\mathrm{N_b}\ \textit{-(e)s}_{\mathrm{Gen.Sg}}]$ — $[\mathrm{N_b}\ \textit{-er}_{\mathrm{Nom.Pl}}]$

$[\textit{Bild}_{\mathrm{N}}\ \textit{-es}_{\mathrm{Gen.Sg}}]$ — $[\textit{Bild}_{\mathrm{N}}\ \textit{-er}_{\mathrm{Nom.Pl}}]$

Why is class membership encoded on the stem? The reason is not that it can always be **predicted** from the stem. The coindices express generalizations over behaviour (same behaviour = same coindex), but generalizations are drawn over fully inflected words (Booij 2010: 2, Masini & Audring 2019, and see Blevins 2006 for arguments that full word forms are the best predictor of inflectional behaviour). Yet, full forms cannot be linked by coindices because they are not in fact the same, they only share the same stem.

This converges with the fact that inflectional class membership is not a property of an individual word form, but a property of the lexeme. Unfortunately, the notion of the lexeme is not easily adoptable in a construction-based theory. Constructions consist of linguistic structure – phonological, morphosyntactic and semantic – while a lexeme is an abstraction and not a piece of structure.[17] Yet, coindexing the stem can go a long way in expressing lexemic relations. Consider example (17). The word forms *go* and *went* belong to the same lexeme, indicated by coindex 1. In the case of *go*, coindex 1 appears in the semantics, the morphosyntax and the phonology, as it would in other forms of the same lexeme, such as *goes* and *going*. In the case of *went*, coindex 1 is absent in the phonology because the phonology is deviant (suppletive). Yet, the form's association with the lexeme GO is vouchsafed by coindex 1 in the semantics and the morphosyntax.[18]

(17)	$[\mathrm{GO}_1]_2$	$[\mathrm{PAST}\ (\mathrm{GO}_1)]_3$
	$[\mathrm{V}_1,\ \textsc{prs}]_2$	$[\mathrm{V}_1,\ \textsc{pst}]_3$
	/gəʊ$_1$/$_2$	/wɛnt/$_3$

In (16) and (17), we see that coindexation of the stem can be used to mark group membership. In the case of individual word forms like *go* and *went*, the group is the lexeme. In (16), the groups are two inflectional classes. Words are often members of various groups at once; for example, *Bilder* is a noun, a member of the $\textit{-(e)s}_{\mathrm{Gen.Sg}}/\textit{-er}_{\mathrm{Nom.Pl}}$ class and a form of the lexeme BILD. This can be encoded by stacked indices, which the theory employs anyway, e.g. in cases where multiple functions converge on one form (see Jackendoff & Audring 2016: 470 for an example).

---

[17] Spencer (2018: 277) puts it concisely: "lexemes are descriptions, not objects".

[18] There are also cases where the coindex appears only in morphosyntax, because the semantics is deviant as well (e.g. in light-verb use, as in *go crazy*). For discussion, see Jackendoff & Audring (in press).

Stem coindexation, then, is a powerful tool for modelling paradigmatic relations. At a closer look, it turns out to be too powerful for modelling paradigms because it is too inclusive. A configuration such as (15) or (16) represents paradigm mates as word forms sharing a stem. However, inflectional paradigms are a special kind of paradigmatic relation, involving more than a shared stem. The German words *Macher* 'maker' and *machbar* 'doable' also share a stem with the verb forms in (15); yet they are not part of the verbal paradigm. Minimally, paradigm mates share the same syntactic category. More importantly, however, they participate in the same feature matrix. In the case of the German verbs, this matrix involves the features person, number and tense, plus mood and voice (which were omitted in (15)). The upshot is that paradigm mates are connected both by shared structure – the stem – and by the fact that the non-shared structure contains information about the same features. The latter fact cannot be stated by sister coindexation alone. The feature matrix requires a mother schema, roughly as in (18). Each inflectional schema and each specific word form has to be a daughter or granddaughter of this schema in order to participate in the paradigm.

(18) [V, TENSE, PERSON, NUMBER, MOOD, VOICE]

The available values for each feature, e.g. singular and plural for number and indicative, subjunctive and imperative for mood, might require additional subschemas. Together with the sister relations sketched in (15), this configuration of schemas can be used to express inflectional paradigms.

An advantage of the network architecture, which permits lower-level information to override higher-level information, is that missing inflectional features can, but need not automatically lead to disqualification for paradigm membership. Inflected words show all manner of deviations from full instantiation of features, both on the level of the individual lexical item (this is usually called *defectiveness*; Sims 2015) or more systematically, e.g. in the neutralization of a feature in the context of another (see e.g. Baerman, Brown & Corbett 2005: 28). In addition, there can be pockets of overdifferentiation, when a word or a group of words shows additional feature values, say a special number or gender value (see e.g. Corbett 2014). As long as the daughter shares enough structure with the mother (and the sisters), it can maintain its place in the paradigm.

Of course, a tantalizing question in this regard is what "enough" means. The theory suggests that override is more forgiveable for some properties than for others (cf. also Booij 2017 on absolute and defeasible properties). For example, diverging in syntactic category instantly disqualifies for paradigm membership, whereas having the wrong stem phonology counts as a mild offence. These are issues that need to be considered in the further development of the theory.

This brief excursion into inflectional morphology reveals an additional functionality of a mother schema: to express the conditions for paradigm membership. Interestingly, this is in essence the same result as in §4.1 and §4.2.2, where we saw the need for a mother. In each case, mother schemas encode generalizations that extend beyond structural overlap between the daughters. The following section draws together the various considerations and offers a brief conclusion.

5. Summary and conclusions

The central question of this paper concerned the way mother nodes and sister relations can divide up the task of encoding morphological knowledge. Starting out with the assumption that a full entry model does not automatically prioritize higher-order generalizations, we inquired about the circumstances under which a mother node is functionally justified. At the same time, the power of sister relations is explored, to see to what extent sameness relations can be encoded by 'horizontal' connections within the constructional network.

Both for same-affix relations and for same-stem relations, sister links turn out to be a powerful modelling tool. Groups of affixed words can be connected by coindexing the affix, and word families can be marked by the coindexation of the shared base. Furthermore, stem coindexation can express a wide variety of paradigmatic relationships, including lexemic relations and inflectional classes. In some cases, e.g. in root-based words such as *splendid* or *ambitious*, sister relations appear to be all that is needed.

In other pairs, such as *childless* and *childish*, pure sister links make less sense, as the relation between them lies in the shared base, which is independently present in the network and can serve as a lexical mother connecting the derived forms. Moreover, a number of specific benefits can be identified for mother nodes that are schematic and contain one or more variables. First, such mother schemas provide the variables needed to state that a pattern is productive. Second, schemas can serve as a relational hub for larger families of complex words, allowing words to connect to the mother instead of having to connect to all sisters. Third, and perhaps most interestingly, mother schemas allow for generalizations that go beyond the structural similarity of the daughters. We have seen three cases:

– different stems with the same function (e.g. *child* and *boy* functioning as the base in *childish* and *boyish*);
– different suffix schemas situated antonymically on a common scale ($[\text{N -}\mathit{ful}]_{\text{A}}$ and $[\text{N -}\mathit{less}]_{\text{A}}$);
– different word forms and their schemas participating in the same inflectional paradigm.

These three cases are united by the fact that they involve relations between structures that are not the same in substance. Such relations resist sister coindexation because they do not involve structural overlap. Rather, they mark non-shared structures as being related via a shared function, a common semantic scale or a feature matrix constituting a paradigm. This yields a clear division of labour between sister links and mother schemas: same-structure relations can be captured by sister links, while other types of relation require a mother to state. Interestingly, then, a lot depends on what information is considered to be encoded in the individual lexical representations of the sisters. The richer a representation, the more structure is available for coindexation. Hence, it is of vital importance that theories are explicit about their understanding of the structure of lexical entries and the information contained therein.

The benefits of mother nodes and sister relations are combined to best effect in second-order schemas. This paper argues that second-order schemas are sister relations between mother schemas. They express the fact that two or more schemas are related because the same lexical material instantiates their variable. Such relations were shown to hold between derivational as well as inflectional schemas, all the way from unproductive [ø *-id*]$_A$ – [ø *-or*]$_N$ (*splendid* – *splendor*), partly productive [N *-ful*]$_A$ – [N *-less*]$_A$ (*careful* – *careless*) to fully productive verbal inflection, as in German [V *-e*$_{1\text{SG.PRS}}$] – [V *-st*$_{2\text{SG.PRS}}$] – [V *-t*$_{3\text{SG.PRS}}$] (*mache* – *machst* – *macht*).

Such configurations are proof of the power of sister links, because they function without a common mother: there is no need for a higher-order schema connecting [ø *-id*]$_A$ and [ø *-or*]$_N$, or [N *-ful*]$_A$ and [N *-less*]$_A$. On the other hand, such sister links are relations between schemas, so they themselves are built from mother nodes. Hence, they represent higher-order generalizations without bloating the taxonomy at the top. Even in a model that exploits rather than avoids redundancy, this appears to be a desirable outcome.

A specific concern of this contribution, then, is to invite further thinking about second-order schemas as a coding means for grammatical knowledge – in morphology, but also in syntax, and in phenomena that straddle the two. Construction-based architectures offer rich opportunities for new modelling solutions in both of these fields.

Acknowledgements

Many thanks to Geert Booij, Martin Hilpert, Steffen Höder, and an anonymous reviewer for their lucid and helpful comments. This paper builds on joint work with Ray Jackendoff, to whom I am deeply indebted and grateful.

References

Baerman, Matthew & Brown, Dunstan & Corbett, Greville G. 2005. *The syntax-morphology interface: A study of syncretism*. Cambridge: Cambridge University Press.

Bauer, Laurie & Lieber, Rochelle & Plag, Ingo. 2013. *The Oxford reference guide to English morphology*. Oxford: Oxford University Press.

Blevins, James P. 2006. Word-based morphology. *Journal of Linguistics* 42(3). 531–573.

Booij, Geert. 2010. *Construction Morphology*. Oxford: Oxford University Press.

Booij, Geert. 2017. Inheritance and motivation in Construction Morphology. In Gisborne, Nikolas & Hippisley, Andrew (eds.), *Defaults in morphological theory*, 18–39. Oxford: Oxford University Press.

Booij, Geert. 2019a. The role of schemas in Construction Morphology. *Word Structure* 12(3). 385–395.

Booij, Geert. 2019b. *The morphology of Dutch*. 2nd edn. Oxford: Oxford University Press.

Booij, Geert & Masini, Francesca. 2015. The role of second order schemas in the construction of complex words. In Bauer, Laurie & Körtvélyessy, Lívia & Štekauer, Pavol (eds.), *Semantics of complex words*, vol. 3, 47–66. Cham: Springer.

Bybee, Joan L. 2010. *Language, usage and cognition*. Cambridge: Cambridge University Press.

Bybee, Joan. 2013. Usage-based theory and exemplar representations of constructions. In Hoffmann, Thomas & Trousdale, Graeme (eds.), *The Oxford handbook of Construction Grammar*, 49–69. Oxford: Oxford University Press.

Cappelle, Bert. 2006. Particle placement and the case for "allostructions". *Constructions*, SV1–7. 1–28.

Cappelle, Bert. 2014. Conventional combinations in pockets of productivity: English resultatives and Dutch ditransitives expressing excess. In Boogaart, Ronny & Colleman, Timothy & Rutten, Gijsbert (eds.), *Extending the scope of Construction Grammar*, 251–292. Berlin: De Gruyter Mouton. (doi:10.1515/9783110366273.251)

Corbett, Greville G. 2014. Higher order exceptionality in inflectional morphology. In Simon, Horst J. & Wiese, Heike (eds.), *Expecting the unexpected: Exceptions in grammar*, 107–126. Berlin: De Gruyter Mouton. (doi:10.1515/9783110219098.107)

Diessel, Holger. 2015. Usage-based construction grammar. In Dąbrowska, Ewa & Divjak, Dagmar (eds.), *Handbook of Cognitive Linguistics*. Berlin: De Gruyter Mouton.

Diessel, Holger. 2016. Frequency and lexical specificity in grammar: A critical review. In Behrens, Heike & Pfänder, Stefan (eds.), *Experience counts: Frequency effects in language*, 209–238. Berlin: De Gruyter.

Diewald, Gabriele. To appear. Paradigms lost – paradigms regained. Paradigms as hyper-constructions. In Sommerer, Lotte & Smirnova, Elena (eds.), *Nodes and links in the network: Open questions in Diachronic Construction Grammar* (Constructional Approaches to Language series). Amsterdam: Benjamins.

Gast, Volker. 2008. Verb-noun compounds in English and German. *Zeitschrift für Anglistik und Amerikanistik* 56(3). 269–282. (doi:10.1515/zaa.2008.56.3.269)

Goldberg, Adele E. 2002. Surface generalizations: An alternative to alternations. *Cognitive Linguistics* 13(4). 327–356. (doi:10.1515/cogl.2002.022)

Hilpert, Martin. 2014. *Construction Grammar and its application to English*. Edinburgh: Edinburgh University Press.

Hilpert, Martin. 2019. Higher-order schemas in morphology: What they are, how they work, and where to find them. *Word Structure* 12(3). 261–273.

Hoffmann, Thomas. 2017. Construction grammars. In Dancygier, Barbara (ed.), *The Cambridge handbook of Cognitive Linguistics*, 310–329. Cambridge: Cambridge University Press.

Hüning, Matthias. 2009. Semantic niches and analogy in word formation: Evidence from contrastive linguistics. *Languages in Contrast* 9(2). 183–201. (doi:10.1075/lic.9.2.01hun)

Jackendoff, Ray S.1975. Morphological and semantic regularities in the lexicon. *Language* 51(3). 639–671. (doi:10.2307/412891)

Jackendoff, Ray. 1997. *The architecture of the language faculty*. Cambridge: MIT Press.

Jackendoff, Ray. 2002. *Foundations of language*. Oxford: Oxford University Press.

Jackendoff, Ray S. & Audring, Jenny. 2016. Morphological schemas. *The Mental Lexicon* 11. 467–493.

Jackendoff, Ray & Audring, Jenny. 2019. Relational Morphology: Morphology in the Parallel Architecture. In Audring, Jenny & Masini, Francesca (eds.), *The Oxford handbook of morphological theory*, 399–408. Oxford: Oxford University Press.

Jackendoff, Ray S. & Audring, Jenny. In press. *The texture of the lexicon*. Oxford: Oxford University Press.

Kapatsinski, Vsevolod. 2013. Conspiring to mean: Experimental and computational evidence for a usage-based harmonic approach to morphophonology. *Language* 89(1). 110–148.

Langacker, Ronald W. 2008. *Cognitive Grammar: A basic introduction*. Oxford: Oxford University Press.

Masini, Francesca & Audring, Jenny. 2019. Construction Morphology. In Audring, Jenny & Masini, Francesca (eds.), *The Oxford handbook of morphological theory*. Oxford: Oxford University Press.

Meesters, Gert. 2004. *Marginale morfologie in het Nederlands*: *Paradigmatische samenstellingen, neo-klassieke composita en splintercomposita*. Gent: Koninklijke Academie voor Nederlandse Taal- en Letterkunde.

Nesset, Tore. 2008. *Abstract phonology in a concrete model: Cognitive linguistics and the morphology-phonology interface*. Berlin: Mouton de Gruyter.

Norde, Muriel & Morris, Caroline. 2018. Derivation without category change: A network-based analysis of diminutive prefixoids in Dutch. In Van Goethem, Kristel & Norde, Muriel & Coussé, Evie & Vanderbauwhede, Gudrun (eds.), *Category change from a constructional perspective*, 47–90. Amsterdam: Benjamins.

Plag, Ingo. 2003. *Word formation in English*. Cambridge: Cambridge University Press.

Sims, Andrea D. 2015. *Inflectional defectiveness*. Cambridge: Cambridge University Press.

Spencer, Andrew. 2018. On lexical entries and lexical representations. In Bonami, Olivier & Boyé, Gilles & Dal, Georgette & Giraudo, Héleène & Namer, Fiammetta (eds.), *The lexeme in descriptive and theoretical morphology*, 277–301. Berlin: Language Science Press. (doi:10.5281/zenodo.1407009)

Van de Velde, Freek. 2014. Degeneracy: The maintenance of constructional networks. In Boogaart, Ronny & Colleman, Timothy & Rutten, Gijsbert (eds.), *Extending the scope of Construction Grammar*, 141–179. Berlin: De Gruyter Mouton.

van der Spuy, Andrew. 2017. Construction Morphology and inflection. *Lingua* 199. 60–71. (doi:10.1016/j.lingua.2017.07.010)

van Marle, Jaap. 1985. *On the paradigmatic dimension of morphological creativity*. Dordrecht: Foris.

Williams, Edwin. 1981. On the notions “lexically related” and “head of a word”. *Linguistic Inquiry* 12(2). 245–274.

Zehentner, Eva & Traugott, Elizabeth Closs. Forthcoming. Constructional networks and the development of benefactive ditransitives in English. In Smirnova, Elena & Sommerer, Lotte (eds.), *Nodes and networks in Diachronic Construction Grammar* (working title). Amsterdam: Benjamins.