Contributions to Chibchan historical linguistics

Pache, M.J.

3 External classification

The present section deals with the external classification of Chibchan. The investigation of hitherto unknown genealogical connections is probably among the most pervasive and intriguing topics in linguistics (cf. Campbell 1997: 26–85, 2003: 262), and with more than 120 language groups and isolates in only Central and South America (Kaufman 1990: 31), this part of the world is among the most thrilling areas for this kind of research.

Evidence for genealogical connections between Native American languages has continuously been discovered since the eighteenth century (e.g., Gumilla 1741: 316; Holt 1975 cited in Campbell 1979: 942–4; Aschmann 1993; Kaufman 1994; Campbell 1995; Adelaar 2000; Viegas Barros 2001, 2005a, 2005b; Carvalho 2009; Ribeiro & van der Voort 2010; Hammarström 2011; Brown et al. 2011, 2014; Seifart & Echeverri 2014; Zamponi 2014; Wichmann 2015; Rosés Labrada 2016; Mora-Marín 2016; Jolkesky 2009, 2017). The external classification of Chibchan has been a topic of inconclusive and intense debate since the late nineteenth century when the existence of the Chibchan language family was recognized by European scholars (cf. Herzog 1886; Uhle 1890), and since then several different genealogical connections of Chibchan with Mesoamerican, Central American, and South American languages have been proposed (e.g., Herzog 1886; Greenberg 1987; Holt 1997/98; Constenla Umaña 2012: 418). In this context of ongoing discussion, the present section aims to identify a plausible candidate, a language or language group that may turn out to be demonstrably related to Chibchan in genealogical terms. What this section will not deal with in detail is language contact of (Proto-)Chibchan.

The structure of this section is as follows: First, I will briefly present some methodological issues of external classification (subsection 3.1). A discussion of previous proposals for the external classification of Chibchan is provided in subsection 3.2. Finally, in subsection 3.3, I present an alternative proposal: evidence for a genealogical connection between Chibchan and Macro-Jê, a language family of eastern South America.

3.1 Analytical framework

The present section will touch upon methodological issues in the investigation of Chibchan external classification. Subsection 3.1.1 will briefly present and discuss some general principles that apply in the identification of non-fortuitous coincidences between languages. Subsection 3.1.2 will elucidate the basic steps that were undertaken in the present thesis to identify the external relations of Chibchan.

3.1.1 The external classification of languages

The process of investigating possible links between languages is hierarchically structured, and the present subsection grossly follows this structure:

- At the initial level, potentially relevant parallels are observed between different languages, and it must be decided whether they are due to contact and/or a genealogical relationship and/or to other origins. This is dealt with in subsection 3.1.1.1.
- If other origins can be ruled out as the only source of observed coincidences, it is relevant to investigate, in a second step, whether the observed parallels originate in contact or cognacy, that is, whether the languages in question share elements because they derive from the same proto-language, because they were in contact with each other, or because of both scenarios. This topic will be dealt with in subsection 3.1.1.2.

For more information about the detection of distant genealogical relationships between languages, see Campbell (e.g., 1997: 206–59, 2003, 2004: 344–60) and Campbell and Poser (2008: 162–223).

3.1.1.1 Assessing the relevance of observed parallels

Procedures that apply in the identification of non-fortuitous relationships between different languages are similar to those undertaken in the reconstruction of the ancestor language of an established language family, briefly discussed in subsection 2.1.1 above. Looking for potentially non-random relationships between two languages, a first step likewise consists in arranging several forms of the languages in question according to their meaning. Imagine we did not know about the genealogical connection that exists between Cabécar, Boruca, Kuna, and Muisca. In this case, we might prepare a table such as table 10, repeated here as table 82.

TABLE 82
ARRANGING SEMANTICALLY RELATED TERMS IN ORDER TO DETECT RECURRENT SOUND CORRESPONDENCES

Cabécar	Boruca	Kuna	Muisca
hĩʃʊkʊ́ ‘earth, floor	*tap* ‘earth, floor’	*nappa* ‘earth, land, ground, floor’	<tybso> ‘potter’s clay’
tɪ ‘cornfield’	*ɾáŋkɾa* ‘sowing, cultivation’		<ta> ‘tillage’
tã́ ‘gourd rattle’	*tá?* ‘gourd rattle’	*naa* ‘gourd rattle’	-/-

Arranging semantically related terms in order to detect recurrent sound correspondences. Sources are Holmer (1952) (Kuna); González de Pérez (1987) (Muisca); Margery Peña (1989) (Cabécar); and Quesada Pacheco and Rojas Chaves (1999) (Boruca).

Likewise, (fossilized) morphology or elements that do not belong to the roots in question should be identified. In the case of the forms shown in table 82, the corresponding roots without remnant morphology are those given in table 11, repeated here as table 83. For more details, see above, subsection 2.1.1.

TABLE 83
SEMANTICALLY RELATED TERMS REFLECTING RECURRENT SOUND CORRESPONDENCES

English	Cabécar	Boruca	Kuna	Muisca
‘earth/clay’	-/-	*tap*	*nap-*	<tyb->
‘field for cultivation’	*tɪ*	-/-		<ta>
‘rattle’	*tã́*	*tá?*	*naa*	-/-

If sound correspondences recur in semantically comparable forms, this may be seen as first evidence for a non-fortuitous relationship between the languages in question. The regular correspondence between the initial consonants of the Chibchan forms given above, in table 83, is illustrated in table 84.

TABLE 84
A REGULAR SOUND CORRESPONDENCE ACROSS DIFFERENT CHIBCHAN LANGUAGES

Cabécar	Boruca	Kuna	Muisca
#t	*#t*	*#n*	<#t>

If several regular sound correspondences can be found between the languages in question, this indicates that the similarities are not the result of chance. It is not necessary to reconstruct the shared ancestor language in order to demonstrate that a genealogical relationship between two languages or language groups exists (cf., e.g., Adelaar 2000; Ribeiro & van der Voort 2010). However, reconstruction can be undertaken once the genealogical connection is made plausible, and if enough cognate sets exist. Indeed, reconstructing proto-forms and checking whether these forms and the postulated sound changes work out helps to further corroborate the existence of a genealogical connection between the languages in question and is part of the comparative method (cf. Campbell 2004: 128–9). Reconstruction is only possible, however, if regular sound correspondences can be detected in a sufficient number of forms in the languages in question.

A priori, coinciding forms can be due to various factors: contact, a shared origin, iconicity, or chance, and differentiating between chance similarities and non-random matches between two languages can be difficult, above all in the case of distant relationships, where the signals of a shared past tend to be faint. The subsequent subsections will touch upon the identification of those parallels that are potentially relevant in the search for external language relations. The following issues will be dealt with: the regularity of sound correspondences, required numbers of potential cognates, sizes of pools from which potential cognates are taken, shared sets of homophones and other forms, shared paradigms, iconicity, semantics of the compared forms and plausibility of the implied scenario.

3.1.1.1.1 Regularity of correspondences

As illustrated in the previous paragraphs, a relevant feature in the identification of a non-fortuitous connection between two languages is the regularity of sound correspondences. Ultimately, it is the regularity of sound change that underlies the regularity of sound correspondences in related languages (Meillet 1925: 29–30) – vice versa, the regularity of sound change is inferred from the observed regularity of sound correspondences between related languages, a somewhat circular aspect of the comparative method. The fact that the relation between the form and meaning of a word is fortuitous (Saussure [1916] 2005: 75) is the second reason why regular sound

correspondences are significant in the detection of non-fortuitous connections between two languages.

The closer two languages are connected, the easier it should be to identify semantically related terms that exhibit systematic, that is, recurrent sound correspondences. The actual similarity of forms, however, is not necessarily relevant. This has been stressed, for instance, by Meillet (1925: 6) in the case of systematic correspondences of Armenian *erk* with a sequence *dw* in Old Greek. An example from Chibchan to be mentioned in this context is the regular correspondence of Muisca <gy> with Kuna *sa*, first proposed by Constenla Umaña (1981), and shown in table 85. Although a priori, the Muisca and Kuna forms in table 85 do not appear very similar, the sound correspondences reflected in the terms in question are regular, and the forms shown in table 85 can be considered as reflecting a non-fortuitous relationship between Muisca and Kuna (which in fact does exist, of course).

TABLE 85
RECURRENT CORRESPONDENCE OF MUISCA <gy> AND KUNA *sa*

Muisca (González de Pérez 1987)	Kuna (Holmer 1952)
<**gy**i> 'mother-in-law of a female ego, daughter-in-law of a male ego' (321)	***sa****kka* 'parent-in-law' (134)
<**gy**e> 'excrement' (261)	*saa* 'posterior, belly, stomach, excrement' (133)

Page numbers in parentheses.

3.1.1.1.2 Number of potential cognates, size of pools

There are probably no absolute benchmarks concerning how many regular sound correspondences are needed or how many coinciding lexical items are required in order to plausibly argue for a genealogical relationship between two languages. For example, four terms of the Patagón de Perico language (Peruvian provinces of Bagua, Amazonas, and of Jaén, Cajamarca) were observed by Rivet (1911: 139) as bearing remarkable similarities with their counterparts in Cariban languages: <tuná> 'water', <anás> 'maize>, <viue> 'firewood' and <coará> 'sheep (sloth)'. These parallels were considered sufficient to demonstrate a Cariban status of Patagón de Perioc by different authors (Torero 1993: 451–2; Adelaar and Muysken 2004: 405–6; Campbell 2012: 83). However, the size of the pools from which coinciding forms have been gathered is important when it comes to the assessment of the (non-)fortuitous character of the observed parallels. In the case of Patagón de Perico, the four words mentioned above

were the entire linguistic material available for this language. Also, twenty lookalikes or words reflecting regular correspondences are a quite robust diagnostic of a systematic relationship if they are taken from a pool of only fifty available comparisons. If the pools contain, say, 5,000 forms for each language, twenty corresponding forms have a much weaker diagnostic value (cf. Rivet 1943b: 78–9). It remains to be established how many regular correspondences are required with relation to the size of the original pools. That is, provided that the absolute number of matching forms is high enough, what kind of ratio of observed correspondences to total number of compared forms is necessary in order to consider observed parallels as non-fortuitous? Four parallel forms which are not too short, like in the case of Patagón de Perico mentioned above, might be the minimum requirement, if the pool of compared elements is small enough.

3.1.1.1.3 Homophonous sets

In the case of a remote genealogical relationship, it might happen that only a few forms have been preserved which reflect systematic sound correspondences, for instance because a lot of analogical changes or lexical replacements have occurred, blurring the picture, or because of sound changes that accumulated and made actual cognate forms so different that they are not easily identified as such (e.g., French *cinq*, English *five*, (Campbell 2004: 349)). Also, formal correspondences can be misleading if they are only found in a few short forms which have a CV structure. In this case, they are more probably the result of chance than if found in coinciding long forms (Campbell 1997: 229). However, this is not the case if these short forms are embedded in sets of shared homophones (Ribeiro & van der Voort 2010: 562–3).[70]

The case of a completely or almost completely homophonous pair that illustrates the systematic connection between two languages (in this case Chibchan) is shown in table 86. Suppose we did not know about any genealogical connection between the Chibchan languages Atanques (northern Colombia) and Boruca (southeastern Costa Rica), and that we observed the terms for ‘wound’, Atanques <sána> and Boruca *ros*, *roská*. This pair alone would not necessarily suggest the existence of a non-fortuitous connection between the two languages. However, as shown in table 86, there is a (nearly) homophonous form, in both languages, with the meaning ‘maize’.

[70] To my knowledge, Ribeiro (2004) is the first scholar to use shared homophones in the demonstration of regular sound correspondences. The only earlier reference to a similar approach is probably Meillet (1925: 6), when he juxtaposes Armenian forms containing the sequence *erk* and Old Greek forms containing the sequence *dw* (see also above, subsection 3.1.1.1.1). However, Meillet did not explicitly mention the concept of shared homophony as demonstrating a non-fortuitous connection between two languages.

TABLE 86
A SHARED PAIR OF (NEAR-)HOMOPHONES IN ATANQUES AND BORUCA

Atanques (Celedón 1892a)	Boruca (Quesada Pacheco & Rojas Chaves 1999)
<sána> 'wound' (596)	*ɾos*, *ɾoská* 'wound' (159)
<sán> 'maize' (596)	*ɾos*, *ɾoská* 'grain' (148)

Page numbers in parentheses.

In table 86, the sound correspondence attested in the first row recurs in the second row, in the context of a semantically unrelated homophone. Provided that sound change is regular (and provided that corresponding terms in the proto-language were homophonous as well), this near homophony reveals the existence of a systematic connection between Atanques and Boruca (cf. Ribeiro & van der Voort 2010: 562–3), which in fact does exist.[71]

Additionally, shared homophones are a relevant diagnostic for a non-fortuitous relationship between two languages in terms of statistical considerations: If two semantically unrelated forms happen to be homophonous in one language (Atanques <sán> 'maize', <sána> 'wound') and in another language (Boruca *ɾos* 'grain', *ɾos* 'wound'), this is difficult to explain as the result of mere chance, especially if there are several homophonous pairs. It goes without saying that from a purely statistical perspective, homophonous triplets have still more diagnostic value in order to demonstrate the existence of a non-fortuitous relationship between two languages than homophonous pairs.

3.1.1.1.4 Other shared sets or paradigms

Another indication of a non-fortuitous connection between two languages are shared paradigms. If, say, two languages A and B share a short number term for 'two', this similarity alone may be dismissed as probably due to chance. Instead, if A and B share sequences of number terms – for instance for 1, 2, 3 or for 7, 8, 9 – an origin of these coincidences in chance can more easily be excluded, given that there are not only parallels concerning single forms, but an entire set, and additionally, position within this

[71] "If, as the comparative method assumes, sound changes tend to be regular, and two different words happen to sound the same in a given proto-language, one expects that these words will still be homophonous in the daughter languages, unless lexical replacement or some irregular factor is at play. Consequently, […] the fact that cases of homophony in one language tend to correspond to homophonous pairs in the other [corroborates] the regularity of the phonological correspondences." (Ribeiro & van der Voort 2010: 562)

set. Sharing several matching sets or paradigms of course, is more robust evidence for a non-fortuitous relationship than the existence of a single shared set or paradigm (cf. Nichols 1997: 360–1; Campbell 2003: 276).

3.1.1.1.5 Iconicity, onomatopoeia, nursery terms

Onomatopoetic forms, nursery terms, and iconic forms including sound symbolism are problematic for the investigation of distant genealogical relationships between languages (Campbell 2003: 272–4). Because of their non-arbitrary form, they should not be considered in the investigation of genealogical relationships (Campbell 2013: 354).[72]

The universal character of nursery ('mama', 'papa') terms was first observed, as it seems, in the eighteenth century (La Condamine 1745: 56–7), and it has been known since this period that if, for instance, Spanish and some indigenous languages of South America share parallels in nursery terms, this is not necessarily the result of recent language contact (Gilij 1782: 414–5), nor of a genealogical relationship.

Other types of sound symbolism may also be problematic when affecting vocabulary used for long-distance genealogical comparisons (Campbell 2003: 273). For instance, sibilants may indicate notions of smallness or affection in several South American languages that are not necessarily closely related, as illustrated by the cases of Proto-Chibchan *^{m}bus- 'young woman' versus *^{m}bu^{n}di 'woman (adult)' (see subsection 2.4.3.2 above), and in (presumably unrelated) Mapudungun *kuʃe* 'old woman (said with affection)' versus *kuðe* 'old woman, pejorative' (Zúñiga 2007: 62). Another example, in the domain of sound symbolism, is nasality. Nasality seems to indicate smallness in different, putatively unrelated South American languages, as in Pumé *ãɾĩ* 'worm' versus *aɾi* 'alligator' (cf. Dyck & Dyck 2015), and in Aguaruna (Jivaroan) *nampitʃ*, 'worm' versus *dapi* 'snake' (cf. Wipio Deicat 2015). If two South American languages share these kinds of alternations in similar contexts, this does not necessarily reflect a particular relationship, but may be due to areal diffusion or more universal principles.[73]

More generally, cross-linguistic sound–meaning correspondences have been shown to affect even the most basic and stable vocabulary (Blasi et al. 2016), so even controlling for the more obvious candidates for sound symbolism does not completely eliminate this sort of confound.

---

[72] Notwithstanding, in some single cases, certain iconic elements may indeed bespeak contact, for instance parallels in exclamatives such as Quechua *alaˈlaw* 'how cold!' (Adelaar & Muysken 2004: 206), Aymara *alaˈlaw*, *alaˈlaj* 'how cold!' (de Lucca 1983: 13), and Cholón *alew* 'how cold!' (Adelaar & Muysken 2004: 464), or the parallel of Quechua *anaˈnaw* (Adelaar 1977: 426), Aymara *anaˈnaj* 'how painful!' (de Lucca 1983: 26) and Mapudungun *ananaj* 'how painful!' (Augusta 1916: 8).

[73] For notions of smallness and affection and their encoding in South American languages see also Regúnaga (2012: 203–4).

3.1.1.1.6 Semantics

Semantics can be a thorny issue in the comparison of languages (Campbell 2003: 271–2). It goes without saying that the elements that are compared across different languages should have the same or at least comparable meanings (cf. table 82 above). It is, of course, not always easy to decide on what counts as comparable in a particular case. Some tendencies seem to hold for many different languages across the world. For instance, a language is more likely to colexify 'bone' and 'leg' rather than 'bone' and 'rib' (cf. Urban 2012: 621). This would justify the comparison of a form *x* 'bone' in language A and a form *y* 'leg' in language B, rather than comparing it with a form *z* 'rib' in language B. However, many semantic associations are rather specific for certain languages, language families or areas (e.g., Urban 2010). Possible examples are the semantic equations 'jaguar/dog', 'leaf/hand' (Adelaar 2013: 124), 'three/alone', and 'four/friend, other' (e.g., Pache 2018b), which occur in different Lowland South American languages. Thus, if a certain sound correspondence is found in the 'leaf' term of language A and in the 'hand' term of language B, this might be a more reliable data point in the comparison of two Lowland South American languages than in the comparison of two languages outside this area.

3.1.1.1.7 Plausibility

In order to be regarded as the result of a non-random origin, parallels between two languages must also be accounted for by a plausible scenario. Formal similarities between, say, English and Totonacan (Wonderly 1953) or Quechuan and Finnish (Campbell 1997: 220) are difficult to bring into line with what is known so far about the prehistory of the groups concerned, and claims that the respective connections are non-fortuitous would require an important amount of additional evidence or correlates from further sources or disciplines such as human genetics or archaeology.

There is also something of a paradox involved in demonstrating a distant genealogical relationship taking into account geography: Related languages (either by contact or by a shared origin) tend to be spoken in geographically close areas (e.g., Wichmann et al. 2010). Therefore, the more distant the areas in which two languages are spoken, the more and better evidence would be needed in order to make a non-fortuitous connection between them plausible. However, the more distant the two areas in question are, the more time tends to have elapsed since the separation of the two ancestor languages in question; and the longer this time, the more changes and lexical replacements may have occurred in the languages in question, blurring the possible evidence for a non-fortuitous connection. To put it differently, in order to demonstrate a non-fortuitous relationship between two languages spoken in distant areas, more evidence would be needed, but less evidence is usually available.

3.1.1.2 Cognacy versus contact

In case a relationship between two languages is arguably non-fortuitous, the question arises whether it is one of contact or of genealogy (or both). Regular sound correspondences between two languages may be found both in borrowings and cognates. Borrowed items, if old, may undergo the same sound changes as words that were inherited (as, for instance, 'hemp' terms in Indoeuropean languages, cf. Lubotsky 1989: 62). However, recurrent sound correspondences which are indicative of a genealogical relationship should be predominantly found in the most stable parts of the lexicon. Within the hundred-words Swadesh list, Holman et al. (2008: 351) identify the following ten terms as the most stable words across the world's languages (in order of decreasing stability): 'louse', 'two', 'water', 'ear', 'to die', 'I', 'liver', 'eye', 'hand', 'to hear'.[74] Stability means that these terms tend to be retained in the history of a language and are relatively rarely replaced by another term, which may either be borrowed or represent a neologism or a word having changed its meaning to the target meaning (ibid.: 334). (This does not preclude the possibility, that in specific cases, these stable terms do become replaced – compare, for instance, Spanish *hígado* 'liver', which is derived from Latin *iecur ficatum* 'fig-stuffed liver' (Campbell 2004: 260).)

If, however, regular sound correspondences are only found in instable parts of the lexicon, including elements that are frequently borrowed, such as certain non-basic kinship terms (Steinen 1886: 288), cultural vocabulary, plant terms (Campbell 2004: 352), or animal terms, there is a strong argument for a contact relationship between two language groups (cf. McMahon et al. 2005). Coming back to an example mentioned above: if two languages happen to share the terms for the numerals 1, 2, 3, this parallel is more probably due to a genealogical relationship than coinciding terms for higher numerals alone, such as 7, 8, 9, which are more prone to being borrowed (cf. Pache 2018b).[75]

If sets of parallels are more indicative of a non-fortuitous connection than single parallels, some have a particularly high diagnostic value for a genealogical relationship: Among these are shared allomorphs or suppletive paradigms of the *good / better / best* or *ist/sind* type (Greenberg 1987: 30). Parallels like these are difficult to explain as the result of language contact or chance, all the more so if several sets of these kinds are

[74] These most stable terms are also at the basis of a lexical database and software developed by Wichmann et al. (2016), the Automated Similarity Judgment Program (ASJP) (see also Holman et al. 2011). It allows for quantifying phonological similarities reflected in 40 stable basic vocabulary items among around two thirds of the world's languages. This powerful instrument is continuously being expanded and revised.

[75] An overview of language contact phenomena in South American languages is provided by Muysken (2012).

shared by two languages. The absence of such parallels, however, is by no means a counterargument to a genealogical connection.

Besides regular sound correspondences in stable parts of the lexicon, it has often been argued that parallels in the morphology of two languages are the most indicative of a genealogical relationship (e.g., Brinton 1891a: 333). Indeed, some grammatical elements, if shared, can be very reliable indicators of a genealogical relationship, such as sets of bound person-marking elements (cf., e.g., Nichols 1997: 361). (A discussion of grammatical evidence in the identification of genealogical relationships can be found in Campbell and Poser 2008: 176–94.) Examples for sets of grammatical elements that have been successfully used in the discovery of genealogical connections are the parallels in person-marking paradigms identified by Lehmann (1910: 720, 1920: 484) in the case of Misumalpan languages, or those identified by Rosés Labrada (2015: 429–63; 2016) in the case of the Sáliba, Piaroa and Mako languages. Other morphological elements, however, are more frequently subject to borrowing (Seifart 2015) and, if shared, are less indicative of a genealogical connection – for instance, classifiers or elements indicating plural number (Seifart 2011).

Finally, it goes without saying that phenomena of contact tend to be attested between two languages that are genealogically related. Genealogically related languages are frequently spoken in geographically close areas, and simultaneously, languages spoken in adjacent or geographically close areas are also more likely to share words via language contact. However, if two languages that are spoken in distant areas share several parallels, it is more difficult to explain these matches by recent contact. In some cases, lexical parallels between geographically very distant languages can be explained by indirect contact, that is, as *Wanderwörter* which diffuse via starburst networks or borrowing chains (cf. Haynie et al. 2014). However, whereas borrowing chains and networks may account for the diffusion of lexical items that refer to salient, culturally important entities or activities – for instance, ritual objects, acculturation terms, and technologies (cf. Haynie et al. 2014) – they may not explain whole sets of coinciding forms that refer to non-salient entities or common activities in geographically distant languages.

If the existence of a genealogical connection between two languages can finally be argued, a major difficulty emerges: the identified genealogical connection between two language groups should be the most obvious (that is, convincing) and the closest possible. Arguing a genealogical connection between, say, French and Dutch is certainly correct, but it is infelicitous if at the same time, the much closer and more obvious relationships between French and Spanish and Dutch and German are overlooked (cf. Greenberg 1987: 4). Thus, the identification of a genealogical link between two languages urges further investigation, and other languages should likewise

be systematically checked for a possible membership in the genealogical unit that is proposed. Another relevant issue is how the languages within the newly established unit relate to each other at an intermediate level (Greenberg 1987: 4). In this sense, proposals of genealogical relationships must be seen as provisional; they call for further calibration, and, if possible, for reconstruction and for subclassification of the languages in question – if more than two are involved – on the basis of shared innovations (for an overview of the comparative method and linguistic reconstruction, see Campbell 2004: 122–47).

3.1.2. Approach in the present study

The present subsection deals with the specific steps that were undertaken in this thesis to identify Chibchan external relations. At all stages, my approach was qualitative.

First, I carried out a reconstruction of Proto-Chibchan lexicon and grammatical elements. Detailed information about Chibchan data and Proto-Chibchan reconstructions has been provided in section 2.

In a second step, the Proto-Chibchan forms (lexicon and grammatical elements) reconstructed in section 2 were used in bilateral comparison with their counterparts in other Native American languages. In line with the principles described above, I particularly focused on two phenomena: First, I searched for homophonous or near-homophonous pairs (two identical/similar forms with different, unrelated meanings) in the non-Chibchan languages, which would correspond to homophones or near-homophones in Chibchan. As stated above, in subsection 3.1.1.1.3, shared homophones may be a convenient heuristic to effectively identify a non-fortuitous connection between two languages. Table 87 shows some examples for Proto-Chibchan (near-)homophones that I considered in this comparison.

TABLE 87
SOME PROTO-CHIBCHAN NEAR-HOMOPHONES

Chibchan (near-)homophones	English
*^{n}du	'father, uncle, ancestor'
*^{n}duʔ	'tooth'
*^{m}ba^{n}d-	'all'
*^{m}ba^{n}d-	'tasty'
*u^{m}ba	'face, eye, fruit'
*(^{n}d-)u^{m}ba	'brother-in-law'

Second, I looked for shared allomorphs (two different forms with an identical/similar meaning). The reason is that shared allomorphy is not very likely to be the result of borrowing or coincidence, especially if it is found in various cases, that is, if several sets of allomorphs are shared between two language groups (see above, subsections 3.1.1.1.4 and 3.1.1.2). Table 88 shows Chibchan allomorphs that I considered.

TABLE 88
SOME PROTO-CHIBCHAN ALLOMORPHS

Chibchan allomorphs	English
*^{m}b-	'to go$_1$'
*^{n}daʔ ~ *taʔ	'to go$_2$'
*a-	valency-reducer$_1$
*aʔ-	valency-reducer$_2$
*a^{n}d-	valency-reducer$_3$
*^{n}d(i)-	relational element$_1$
*ʦ(i)-	relational element$_2$

In my bilateral comparisons of Proto-Chibchan with other Native American languages, I focused on languages from eastern South America for different reasons, non-linguistic and linguistic. Among the non-linguistic reasons were the following two:

1. Human genetics suggest some similarities of Chibchan speaking groups and populations of eastern South America in the domain of autosomal markers (see above, subsection 1.4.1).
2. Archaeology suggests the existence of a particular link between populations of the Intermediate Area with eastern South America in cultural terms, for instance in the domain of stonework: there are specific coincidences between artifacts (split greenstone pendants) from Costa Rica, from the Tairona area in northern Colombia, and from the Trombetas area, on the northern side of the lower Amazon (see above, subsection 1.4.2).

The linguistic reasons to hypothesize genealogical connections of Proto-Chibchan in South America, especially eastern South America, were the following:

1. The highest number of Chibchan connections suggested by Herzog (1886), and by Rivet (1924a), Greenberg (1987) and other adherents of Macro-Chibchan have been proposed for (apparently very diverse) South American languages (e.g., Cariban, Barbacoan, Páez), but much fewer for languages of Meso- or North America. This is remarkable, even if Macro-Chibchan proposals seem to be frequently based on lookalikes rather than on systematic coincidences.
2. Several body-part terms in Chibchan languages (for instance, 'eye', 'ear', 'mouth') are morphologically complex (see subsection 2.2 above). Analyzability of body-part terms is a relatively widespread phenomenon in South American languages, but much less so in North American languages (Urban 2012: 184–5).
3. Reflexes of an element *k^{w}a 'seed' and *ka^{n}d- ~ *kat- 'stick, bone, tree' are frequently found as stem formatives in Chibchan body-part terms (cf. Constenla Umaña 1988, 2012: 405–7, 412; Pache 2016a), for instance in Chimila /kak**kwa**/ 'mouth' and /hatta**kra**ʔ/ 'hand' (cf. Malone 2005: 202). Originally, the element *k^{w}a may have referred to small or to roundish entities, irrespectively of their concavity or convexity. Similar elements are attested in Proto-Jê *j-ar**kua** 'mouth', and also in Leko *kul**wa*** 'inside of ear' (cf. *kul* 'ear') (Leko data are from Simon van de Kerke, n.d.). Proto-Chibchan *ka^{n}d- ~ *kat- 'stick, bone, tree', reflected as a stem formative in Chimila /hatta**kra**ʔ/ 'hand', is remarkably similar to the Nambikwaran (central dialect) classifying elements <-kere> for hard, rigid entities and <-kate> for longish entities (cf. Lévi-Strauss 1948: 186), and also to Qawasqar *qar* 'tree, shell, bone' and *q^{h}at* 'foot, leg' (Pache 2016a). As its lookalikes in Chibchan languages, Qawasqar *qar* is frequently attested in body-part vocabulary such as *je**rqar*** 'neck', *qjep**qar*** 'chest', and *qjot**qar*** 'back' (cf. Clairis and Viegas Barros 2015).

4. Nasality is phonemic in vowels and determines the nasality of preceding voiced stops in Bribri, Cabécar, and probably also in Bocotá, Chibchan languages of Central America (see subsection 2.3 above). Vowel nasality prevailing over consonant nasality was most probably also a feature of Proto-Chibchan (cf. Constenla Umaña 1981, 2012; see also above, subsection 2.4.1.4), and this phenomenon is attested in several Lowland South American languages: Tupían, Macro-Jê and Tucanoan languages, among others. This parallel between the latter language groups and Chibchan was first explicitly noticed, to my knowledge, by Constenla Umaña (1985a: 375). Also, the existence of nasality contrasts in vowels but not in consonants, as it probably existed in Proto-Chibchan, is more frequently attested in South American languages than in North American languages (cf. Maddieson et al. 2014, cited in Chevrier 2017a: 145).

Table 89 shows the North and Mesoamerican languages that have been included in the quest for Chibchan external relations.

TABLE 89
NORTH AND MESOAMERICAN LANGUAGES THAT HAVE BEEN COMPARED WITH CHIBCHAN AND PROTO-CHIBCHAN

	Language or language group	Source
1	Algic: Proto-Algic	Proulx 1984
2	Bella Coola or Nuxalk (Salishan)	Key 2015a
3	Chatino, Zacatepec variety (Oto–Manguean)	Pride and Pride 2015
4	Coahuilteco (unclassified)	Swanton 1940
5	Comecrudo (unclassified)	Swanton 1940
6	Cotoname (unclassified)	Swanton 1940
7	Cuitlatec	Escalante Hernández 1962
8	Karok or Karuk (unclassified)	Bright 2015
9	Huave: Proto-Huave	Noyer 2012
10	Mixe–Zoquean: Proto-Mixe–Zoquean	Wichmann 1995
11	Purépecha (unclassified)	de Wolf 1991; Chamoreau 1998
12	Seri (unclassified)	Moser & Marlett 2015

13	Tequistlatecan: Highland Chontal	Turner 1966
14	Timucua (unclassified)	Granberry 1993
15	Totonacan: Proto-Totonacan	Brown et al. 2011
16	Tsimshian	Key 2015b
17	Uto–Aztecan: Proto-Uto–Aztecan	Voegelin et al. 1962; Miller 1967; Stubbs 2011
18	Waikuri (unclassified)	Zamponi 2004
19	Zuñi (unclassified)	Walker 2015

Table 90 shows the Central American languages that have been compared with Chibchan here.

TABLE 90
CENTRAL AMERICAN LANGUAGES THAT HAVE BEEN COMPARED WITH CHIBCHAN AND PROTO-CHIBCHAN

	Language or language group	Source
1	Jicaquean: Proto-Jicaquean (Proto-Tol), Jicaque (Tol)	Campbell & Oltrogge 1980; Holt 1999b
2	Lencan: Proto-Lencan, Honduran Lenca, Salvadorian Lenca	Hernández & Pinart 1897; Lehmann 1920; Arguedas Cortés 1988
3	Misumalpan: Proto-Misumalpan, Matagalpa, Miskito, Ulwa	Lehmann 1920; Constenla Umaña 1987, 2005; Green 1999
4	Mayan: Proto-Mayan	Kaufman 2003
5	Xinkan: 18th-century Xinka	Sachse 2010

Finally, table 91 shows the South American languages that have been compared with Chibchan for this thesis.

TABLE 91
SOUTH AMERICAN LANGUAGES THAT HAVE BEEN COMPARED WITH CHIBCHAN AND PROTO-CHIBCHAN

	Language or language group	Source
1	Andakí (unclassified, possibly related to Páez)	Rivet 1924b, Gómez & Torres 2012/13
2	Andoque (unclassified)	Landaburu 2000b
3	Aikanã (unclassified)	Becker-Donner 1955; van der Voort 2005
4	Arara do Rio Branco (unclassified)	D'Angelis 2010
5	Arawá: Proto-Arawá	Dixon 2004
6	Aticúm (unclassified)	Meader 1978
7	Awaké (unclassified)	Migliazza 1978
8	Barbacoan: Proto-Barbacoan, Awa, Cha'palaachi or Cayapa, Tsafiki or Colorado, Totoró	Huber & Reed 1992; Curnow and Liddicoat 1998; Moore 2015; Wiebe & Wiebe 2015
9	Bora–Witotoan: Proto-Witotoan, Bora, Miraña, Muinane, Minika, Murui, Nüpode, Ocaina	Huber & Reed 1992; Aschmann 1993
10	Boróro (not necessarily Macro-Jê according to Jolkesky 2016: 264–74), not considered as a Macro-Jê language by Nikulin & Carvalho 2018).	Crowell 1979
11	Cahuapanan: Proto-Cahuapanan, Chayahuita	Hart 1998; Valenzuela Bismarck 2011
12	Canichana (unclassified)	Créqui-Montfort & Rivet 1913; Crevels 2012
13	Cariban: Proto-Cariban, Proto-Taranoan, De'kwana, Macushi, Panare, Tiriyó/Trio; Wai Wai	Meira 1998, 1999; Carlin 2004; Gildea & Payne 2007; Abbott & Foster 2015; Hall 2015; Hawkins 2015; Price 2015
14	Cayuvava (unclassified)	Key 2015c
15	Chapacuran: Proto-Chapacuran, Pacaas Novos	Angenot-de Lima 1997; Misión Nuevas Tribus 2015
16	Chocoan: Proto-Chocoan, Emberá, Epena Pedee (Saija), Waunana	Constenla Umaña & Margery Peña 1991; Holmer 1963; Pardo Rojas 2015; Quiro Dura & Harms 2015
17	Cholón–Hibito: Cholón	Alexander-Bakkerus 2005

18	Chon: Gününa Küne (Puelche), Tehuelche	Delahaye 2015; Viegas Barros & Casamiquela 2015
19	Chono (unclassified)	Bausani 1975
20	Cofán (unclassified)	Borman 2015
21	Culli	Martínez Compañón 1782–90
22	Esmeraldeño (unclassified)	Seler 1902
23	Fulniô, Yatê (unclassified; connected to Macro-Jê, e.g., by Rodrigues 1999)	Lapenda 1965
24	Gamella (unclassified)	Nimuendajú 1937
25	Guahiboan, Proto-Guahiboan	Christian & Matteson 1972; Queixalós 1993
26	Guaicuruan: Mocoví, Pilagá, Toba	Buckwalter & Ruiz 2015; Buckwalter & Suárez 2015; Buckwalter & Sánchez 2015
27	Guamo (unclassified)	Anonymous [1788] 1928
28	Guató (connected to Macro-Jê, e.g., by Rodrigues 1999)	Palácio 1984; Postigo 2009
29	Harakmbut–Katukinan	Adelaar 2000
30	Huarpean: Allentiac, Millcayac	Tornello et al. 2011
31	Irántxe or Münkü (unclassified)	Monserrat 2010
32	Itonama (unclassified)	Key 2015d
33	Jaqi: Aymara	De Lucca 1983; Layme Pairumani 2015; Emlen 2017
34	Jê: Proto-Jê; Apinajé, Canela Apãniêkrá, Mẽbêngôkre, Panará; Xaváante; Proto-Southern Jê, Kaingáng	Davis 1966; Lachnitt 1988; Dourado 2001; Reis Silva 2003; Castro Alves 2004; Oliveira 2005; Jolkesky 2010; Ribeiro & van der Voort 2010; Key 2015e, 2015f; Nikulin 2015b, 2017
35	Jirajaran: Ayomán, Gayón, Jirajara	Jahn 1927
36	Jivaroan: Aguaruna	Wipio Deicat 2015
37	Kakua–Nukak: Kakua, Nukak	Huber & Reed 1992; Bolaños 2016
38	Kamakã (not necessarily Macro-Jê according to Jolkesky 2016: 274–84)	Martins 2007
39	Kamsá (unclassified)	Huber & Reed 1992
40	Kandoshi–Shapra (unclassified)	Tessmann 1930

41	Kanoê (unclassified)	Bacelar 2004; van der Voort 2005
42	Karirí (probably Macro-Jê; not necessarily Macro-Jê according to Jolkesky 2016: 264–74)	Adam 1897
43	Krenák (Botocudo, Borum) (Macro-Jê according to Nikulin 2018; not necessarily Macro-Jê according to Jolkesky 2016: 274–84)	Pessoa 2012
44	Kunza (unclassified)	Vilte Vilte n.d.; Lehnert Santander 2015
45	Kwaza (unclassified)	Van der Voort 2004, 2005
46	Leko (unclassified)	Van de Kerke n.d., 1998, 2000, 2002, 2006a, 2006b, 2009
47	Lengua–Maskoy: Sanapaná (Enlhet dialect)	Unruh & Kalisch 2015
48	Lule–Vilelan: Lule, Vilela	Machoni de Cerdeña 1732; Gilij 1782
49	Macro-Jê: Proto-Macro-Jê, Proto-Jê; Karajá, Chiquitano, Proto-Jabutí	Davis 1966, 1968; Pires 1992; Adelaar 2008; Ribeiro & van der Voort 2010; Ribeiro 2012; Key 2015g; Nikulin 2015a
50	Maipuran (Arawak)	Payne 1991; Aikhenvald 1999
51	Máku (unclassified)	Migliazza 1978
52	Mapudungun (unclassified)	Augusta 1916; Smeets 2008
53	Matacoan: Chorote, Wichí	Gerzenstein 1999, 2015; Braunstein 2015
54	Matanawí (unclassified)	Nimuendajú 1925
55	Maxakalí (Macro-Jê according to Nikulin 2018; not necessarily Macro-Jê according to Jolkesky 2016: 274–84)	Campos 2009
56	Mochica (unclassified)	Middendorf 1892
57	Mosetén–Chimané	Gill 1999; Sakel 2004; Pérez Diez 2015
58	Movima (unclassified)	Haude 2006; Judy 2015
59	Munichí (unclassified)	Tessmann 1930; Gibson 1996
60	Mura–Pirahã	Nimuendajú 1925

61	Nadahup: Proto-eastern Makú; Hup	Martins 2005; Epps 2008b
62	Nambikwaran: Sabanê, Maimandê	Araujo 2004; Eberhard 2009
63	Natú (unclassified)	Meader 1978
64	Ofayé (Macro-Jê according to Nikulin 2018; not necessarily Macro-Jê according to Jolkesky 2016: 264–74)	Oliveira 2006
65	Omurano (unclassified)	Tessmann 1930
66	Otí or Xavante (unclassified)	Quadros 1892
67	Otomacoan: Otomaco, Taparita	Rosenblatt [1936] 1964
68	Páez (unclassified)	Rojas Curieux 1998; Jung 2008; Gerdel 2015
69	Pankararú (unclassified)	Loukotka 1963
70	Pano–Tacanan: Proto-Panoan, Proto-Tacanan	Shell 1965; Girard 1971
71	Peba–Yagua: Yagua	Powlison & Powlison 2015
72	Puinave (unclassified, possibly related to Kakua–Nukak)	Girón Higuita 2008; Richardson 2015
73	Pumé (Yaruro) (probably related to Chocoan)	Mosonyi, Mosonyi & Ramón García 2000; Dyck & Dyck 2015
74	Puquina (unclassified)	Torero 2002
75	Purí (connected to Macro-Jê, e.g., by Rodrigues 1999)	Neto 2007
76	Qawasqar (unclassified)	Clairis 1985; Clairis & Viegas Barros 2015
77	Quechuan: Tarma Quechua, Bolivian Quechua	Adelaar 1977; Rosat Pontacti 2004; Emlen 2017
78	Rikbaktsá (Macro-Jê according to Nikulin 2018; not necessarily Macro-Jê according to Jolkesky 2016: 264–74)	Silva 2011
79	Sáliba–Piaroan: Betoi, Mako, Piaroa, Sáliba, Yuwana (Jodï)	Estrada Ramírez 2000; Mosonyi 2000; Zamponi 2003; Rodman & Rodman 2015; Rosés Labrada 2015
80	Sapé or Kaliana (unclassified)	Migliazza 1978
81	Sechuran–Catacaoan	Urban 2015
82	Tallán (unclassified)	Martínez Compañón 1782–90

83	Taruma (unclassified)	Loukotka 1949; Meira 2015
84	Taushiro (unclassified)	Alicea Ortiz 1975a, 1975b
85	Tequiraca (Auischiri) (unclassified)	Tessmann 1930
86	Ticuna–Yurí–Carabayo	Tessmann 1930; Carvalho 2009; Seifart & Echeverri 2014
87	Timote–Cuica: Timote	Jahn 1927
88	Tinigua (unclassified)	Ernst 1891
89	Trumai (unclassified)	Guirardello 1999; Monod-Becquelin 2015
90	Tucanoan: Proto-Tucanoan	Waltz & Wheeler 1972
91	Tupian: Proto-Tupían, Mekéns, Tuparí; Proto-Mawetí–Guaraní, Awetí, Proto-Tupí–Guaraní, Guaraní	Borella 2000; Mello 2000; Galúcio 2001; Alves 2004; Silva 2007; Rodrigues 2007; Rodrigues & Cabral 2012; Guasch & Ortiz 1991; Céspedes & Canese 2015; Meira & Drude 2015
92	Tuxá (unclassified)	Meader 1978
93	Umbra (unclassified)	Rendón 2011
94	Urarina (unclassified)	Tessmann 1930; Olawsky 2006
95	Uru–Chipayan: Chipaya, Ch'imor Uru, Uchumataqu	Olson & Olson 2015; Cerrón-Palomino 2006, 2016; Hannß 2008
96	Waorani (unclassified)	Peeke 2015
97	Warao (unclassified)	Mosonyi, Mosonyi & Arintero 2000
98	Xukuru (unclassified)	Meader 1978
99	Yagán (Yámana) (unclassified)	Guerra Eissmann 2015
100	Yuracaré (unclassified)	Van Gijn 2006
101	Yurumanguí (unclassified)	Ortiz 1946
102	Zamucoan: Ayoreo, Chamacoco	Ciucci 2013; Briggs 2015
103	Zaparoan: Iquito, Záparo	Tessmann 1930

In case the bilateral comparison of data from Proto-Chibchan with data from other Native American languages (single languages or proto-languages, when reconstructions were available) brought about promising results, I additionally undertook comparisons with data from all single Chibchan languages listed in table 1 (subsection 1.1.1) with the Non-Chibchan language or language group in question. The underlying assumption was

that in some cases, single languages may retain specific cognate elements which do not appear in the reconstructed proto-languages.

Also, I discuss several proposals for connections that have been repeatedly suggested in previous publications, such as Chibchan–Misumalpan, Chibchan–Lencan, or Chibchan–Páez (e.g., by Rivet 1924a; Greenberg 1987; Constenla Umaña 2012). This topic is dealt with in subsection 3.2. The problematic issue which arises in the discussion of previous proposals is the following: Assessing them, one cannot prove that languages are *not* genealogically related. As a strategy of sorts, one can only shift the focus, if possible, to more probable cases by suggesting alternative and more obvious connections based on parallels, say, in basic vocabulary. This approach may be illustrated by the following example: Several authors have claimed the existence of a close genealogical relationship between Chibchan and Misumalpan languages. However, as will be argued below, the evidence presented so far for a genealogical connection between Misumalpan and Chibchan turns out *not* to be more convincing than the evidence for an alternative genealogical connection of Misumalpan with a non-Chibchan language, Páez. This observation casts doubt on the hypothesis that Chibchan is the best candidate for a close genealogical relationship with Misumalpan (see below, subsection 3.2.2.1). However, it does not exclude the possibility that Misumalpan and Chibchan are more distantly related (see also above, subsection 3.1.1.2).

3.2 External classification of Chibchan: Previous proposals

In this subsection, I will briefly discuss several external relations of Chibchan that have been proposed so far. I will not elaborate on putative genealogical links of Chibchan with Uto–Aztecan or Barbacoan which have been critically assessed upon several occasions by Quesada (2007: 41–2, for Barbacoan; the Chibchan–Barbacoan connection was also refuted by Holmer 1947: 206) and Constenla Umaña (1981: 8–11, 63–4, 2012: 418, for Barbacoan and Uto–Aztecan).

3.2.1 North- and Mesoamerican languages

Relatively few proposals of genealogical or contact relationships have been put forward for Chibchan and Mesoamerican or North American languages. Jijón y Caamaño (1943: 466) saw an affinity ("afinidad") of an alleged Macro-Chibchan phylum (containing, among other languages, Páez and Barbacoan) with Subtiaba, an Oto–Manguean (Tlapanec–Manguean) language, and with Hokan, a hypothetical language family of western North America – that is, Jijón y Caamaño (1943) added Macro-Chibchan to Lehmann's (1920: 973) and Sapir's (1925) putative Subtiaba–Hokan. Swadesh (1954) and Holt (1997/98) suggested the existence of a different genealogical connection, namely between Chibchan and Uto–Aztecan. The latter proposal has been negatively assessed by Constenla Umaña on different occasions (1981: 63–4; 2012: 418). Greenberg (1987: 106) argued for a relatively close genealogical connection between Chibchan and Purépecha, which is likewise dismissed by Constenla Umaña (2012: 418) and Bellamy (2018: 88). Holmer (1947) pointed to some similarities of Chibchan languages with Timucua (an unclassified, extinct language of northern and central Florida and southeastern Georgia). The last two connections, Chibchan–Purépecha and Chibchan–Timucua will be presented here in more detail.

3.2.1.1 Chibchan–Purépecha

Purépecha is an unclassified language which was spoken in a large part of what today is the State of Michoacán in western Mexico, when European invaders arrived in this area. At the beginning of the 1990's, there were still almost 100,000 speakers of this language (Chamoreau 1998: 32–4). Among the languages of Mesoamerica, Purépecha stands out for its unusual typological profile (cf. Campbell et al. 1986): the only morphological device of Purépecha is suffixation (apart from some cases of reduplication), and accusative case is indicated by a clitic *=ni* (Foster 1966). It has been impossible, so far, to connect Purépecha with any other known language of Mesoamerica. Some authors have linked it with Quechuan (e.g., Swadesh 1967; Liedtke

1997: 75–86), and Anawalt (1992) observed similarities in the garment styles of the Purépecha and of indigenous populations of Ecuador. Greenberg (1987: 106) instead considered Purépecha to be relatively closely related to Chibchan. Among the parallels of Purépecha with Chibchan unravelled by Greenberg (1987) is a final element *-kua* with several functions in Purépecha, among which is nominalization (cf. Chamoreau 2000: 311–2). Both in form and function, Purépecha *-kua* resembles Proto-Chibchan *k^{w}a 'seed, fruit' and its reflexes as a derivational element in Chibchan languages (cf. Constenla Umaña 1988; Pache 2016a; and subsection 3.3.4.2 below). Nevertheless, a detailed assessment of the linguistic evidence, prepared by Bellamy (2018), suggests that none of the genealogical connections proposed so far with South American languages (Quechuan, Chibchan) can be sufficiently substantiated so as to constitute serious claims.

3.2.1.2 Chibchan–Timucua

Timucua is an extinct, unclassified language formerly spoken in what corresponds to present-day northern Florida and southeastern Georgia (Granberry 1991: 195). Holmer (1947: 212–7) was probably the first author to observe some similarities of Kuna and Kogi with Timucua. He attributed them to contact across the Caribbean Sea (ibid.: 217). In terms of lexical parallels, Holmer mentions, for instance, Timucua *abo* 'maize' versus Kuna *opa* 'maize' (a *Wanderwort*, widespread through Central and South America, according to Lyle Campbell, p.c.), and Timucua *uli* 'pot' versus Kuna *ulu* 'pot' (ibid.: 216–7). To this list, one might add Timucua <yaba> 'medicine man' (Granberry 1993: 212) versus Gualaca (Dorasque) <yápa> 'priest' (Pinart 1890: 29) (possibly to be linked with Proto-Chibchan *^{n}da^{m}ba 'feline$_1$'). Also, Holmer (1947: 215) observes that a common element in Timucua is a suffix *-k^{w}a*, which is central in Chibchan word-formation (and in Purépecha, see above). It hardly occurs in Maipuran languages from South America but is more frequent among the Maipuran languages once spoken on the Caribbean islands, according to Holmer (ibid.).

Greenberg (1987: 107) classifies Timucua as a Paezan language, within his Chibchan–Paezan subgroup of Amerind. Granberry instead (1991: 204) sees Timucua as a creole language, the predominant grammatical contributors being Kuna and, above all, Warao, an unclassified language of the Orinoco Delta in Venezuela (likewise, Warao is classified as Chibchan–Paezan by Greenberg 1987: 106). Further parallels of Timucua are identified by Granberry (1991: 216–20) in some South American languages: Chocoan, certain Maipuran languages and Tucanoan. Based on that, he argues for a homeland of Timucua in northwestern Amazonia (Granberry 1993: 41).

I found no indications for a close genealogical connection between Chibchan and Timucua. Nonetheless, a South American origin of the Timucua language would,

indeed, explain the parallels with Chibchan, mentioned above, and the existence of some parallels between Timucua and Leko, a nearly extinct, unclassified language of the eastern foothills of the Andes, near Lake Titicaca. So far, this language has been investigated above all by van de Kerke (n.d., 1998, 2000, 2006a, 2006b). Both Timucua and Leko share a small paradigm of elements indicating person, presented in table 92.

TABLE 92
TIMUCUA–LEKO LOOKALIKES IN THE DOMAIN OF PERSON MARKING

	Timucua	Leko
1	<-na> 'my' (206)	***eɾa*** 'I' (2000: 27)
2	<-ya, -ye> 'your' (232)	***ija*** 'thou' (2000: 27), *jV-* 'me' (2000: 37)
3	<he-ca, ni-heca> 'we' (84)	*heka* 'you (pl.)' (2000: 27)[76]

Sources are Granberry (1993) (Timucua) and van de Kerke (2000) (Leko). Page numbers in parentheses.

More coinciding elements, including body-part terms (sets 4 to 7), and other forms (sets 8 to 17), are presented in table 93.

TABLE 93
SOME FURTHER TIMUCUA–LEKO LOOKALIKES

	Timucua	Leko
4	<chini> 'nose, nostrils' (214)	***ʧin****wa* 'nose, tip' (2009: 309)
5	<sili> 'forehead, tomato, brains, senses' (167)	*siɾi* 'eye' (2009: 309)
6	<hue> 'hand, fingernail' (137)	*wi* 'finger' (n.d.)
7	<**pena**ni> 'with hands' (204)	*pel* 'arm' (2009: 290)
8	<hibi> 'to capture' (188); <**ebe**ta> 'to catch' (225)	***hep****ka-* 'to grab' (2002: 250)
9	<hibe> 'louse' (211)	***hepe****l* 'tick; a small mosquito (Sp. *polvorín*)' (van de Kerke n.d.)

[76] This parallel would be in line with an originally egophoric system in a hypothetical Proto-Timucua–Leko.

10	<**oba**cha> ‘kiss’ (208); <hoba, huba, homa> ‘love’ (211) [unclear whether noun or verb]	***opa****sif* ‘to kiss’ (n.d.)
11	<hiba> ‘to sit, dwell, remain, live’ (135)	*hep*^{h}*o*, *hebo* ‘bird’s nest’ (2009: 312)
12	<heba> ‘to speak’ (228)	*eba-ʧ* ‘to say’ (2000: 28)
13	<ela> ‘sun’ (230)	*hena* ‘sun’ (n.d.)
14	<to> ‘to put’ (220)	*to-* ‘to place, put’ (n.d.)
15	<qili> ‘combat’ (190)	*kel-* ‘to hit’ (2002: 201)
16	<mine> ‘one’ (216)	*ber* ‘one’ (1998: 200)
17	<tafi> ‘brother-in-law’ (187), ‘sister-in-law’ (227)	*ʧab* ‘brother-in-law’

Sources are Granberry (1993) (Timucua) and van de Kerke (1998, 2000, 2002, 2006a, 2009, n.d.) (Leko). Page numbers in parentheses.

I hasten to add that these sets of Timucua–Leko lookalikes are only preliminary, and that remnant morphology in the Timucua and Leko forms needs to be thoroughly analyzed in future studies. Also, the semantics in the lookalikes presented in table 93 do not always fit perfectly, as in (11) ‘to sit’/’nest’ (note, however, that English ‘nest’ likewise contains a reflex of Proto-Indo-European *sed- ‘to sit down’, see Wodtko et al. (2008: 590–91)). Notwithstanding, some vowel correspondences reflected in the nineteen forms given in tables 92 and 93 seem to recur, and it seems difficult to argue that they are purely fortuitous. The vowel correspondences in question are summarized in table 94.

TABLE 94
TIMUCUA–LEKO VOWEL CORRESPONDENCES FROM TABLES 92 AND 93

Timucua	Leko	Set
<a>	*a*	1, 2, 3, 10, 12, 13, 17
	o	11
<e>	*e*	3, 7, 8, 9, 12, 13
	i	6
<i, y>	*i, j*	2, 4, 5, 15
	e	8, 9, 11, 15, 16
<o>	*o*	10, 14

Although the data presented here only allow for hypotheses, the recurrent vowel correspondences shown in table 94 are remarkable insofar as they are attested in a set of three person-marking elements and Leko and Timucua body-part terms for 'forehead/eye', 'hand/finger', 'hand/arm', and 'nose'. The latter terms are among the most stable lexical items across the languages of the world (cf. Holman et al. 2008). Notwithstanding, the Timucua data cited in tables 92 and 93 are not from primary sources on this language, and in the Timucua dictionary edited by Granberry (1993), there are often several divergent Timucua forms for the same English entry. For the latter observation, scenarios such as language mixing might also eventually be envisaged; the matter requires further investigation. Although a non-fortuitous Timucua–Leko connection would be difficult to believe for the sheer distance between both areas, the South American homeland of Timucua implied by the lookalikes presented here is perfectly in line with Granberry's (1993) proposal of Amazonian connections of Timucua (see above), and with Holmer's (1947) observation of language contact between Timucua and Chibchan.

3.2.2 Central American languages

Several authors have argued the existence of genealogical connections of Chibchan with more or less adjacent indigenous languages of Central America, such as those of the Misumalpan family (e.g., Greenberg 1987: 106; Constenla Umaña 2005, 2012: 418), Xinkan (Greenberg 1987: 106), and Lencan (Greenberg 1987: 106; Constenla Umaña 2012: 418). Although Campbell (1997: 79) does not favor a classification of Chibchan, Misumalpan, Xinkan, and Lencan as Macro-Chibchan, at present, the alleged genealogical relationship with Misumalpan and Lencan seems to be considered the most promising external relation of Chibchan (Constenla Umaña 2012: 418). However, as will be argued here, the evidence of alternative connections of Misumalpan and Lencan is not less convincing than the evidence provided so far for a connection of these languages with Chibchan.

3.2.2.1 Chibchan–Misumalpan

Misumalpan languages were once spoken in an area extending from (north-)eastern and central Nicaragua to eastern San Salvador. The family consists of three subgroups, Miskito, Sumu, and Cacaopera-Matagalpa. Extinct Cacaopera was spoken in the Morazán department of northeastern El Salvador, relatively far away from its closest relative, Matagalpa. Matagalpa, likewise extinct, was formerly spoken in the central highlands of Nicaragua and in adjacent areas of southern Honduras (Lehmann 1920:

479; Campbell 1975; Craig 1985; Campbell 1997: 167). Within the Isthmo-Colombian area, the presence of Matagalpa toponyms in Chinandega (cf. Constenla Umaña 1994: 195) indicates that the northern Pacific coast of Nicaragua was populated by Misumalpan-speaking groups before the arrival of Mesoamerican populations from the north, such as the Chorotega (speaking a Manguean language) and the Nicarao (speaking a variety of Nahua, a Uto–Aztecan language) (Constenla Umaña 2002: 190).

At present, Miskito and Sumu (the latter being actually composed of two languages: Mayangna and Ulwa) are the only surviving members of the family. Mayangna and Ulwa are not mutually intelligible (Green 1999: 10–1). Miskito has some 180,000 speakers in northeastern Nicaragua (North Caribbean Coast Autonomous Region) and adjacent parts of Honduras, according to the Ethnologue catalogue (Simons & Fennig 2017). Mayangna is still spoken by some 10,000 people in northern Nicaragua and parts of Honduras (Benedicto et al. 2007), whereas Ulwa is only spoken by some 350 adults in Karawala, a village on the Nicaraguan Atlantic coast (South Caribbean Coast Autonomous Region) (Koontz-Garboden 2009).

The term ‘Misumalpan’ was coined by Mason (1940: 75) (Johnson 1940: 110 used the term ‘Mosumalpan’). Lehmann (1910: 719–22) was the first researcher to identify the genealogical unit of Misumalpan and to propose three subgroups consisting of Miskito, Sumu and Matagalpa-Cacaopera, respectively. This genealogical unit is most obvious in parallels of these languages’ verbal and nominal person marking paradigms (cf., e.g., Lehmann 1920: 484). Reconstruction of Proto-Misumalpan lexicon is often hampered by the fact that the lexicon of Miskito is not necessarily very similar to the lexicon of its sister languages Sumu and Cacaopera-Matagalpa. In line with that, Miskito has been classified as relatively isolated within the Misumalpan family (cf. Campbell 1997: 167).

A connection of Misumalpan and Chibchan groups has frequently been hypothesized (e.g., Lehmann 1920: 645; Mason 1950: 174; Loukotka 1968: 250–3). Greenberg (1987: 106, 382) classifies Chibchan and Misumalpan languages as most closely related – both would belong to “nuclear Chibchan” according to his classification. Constenla Umaña (2002, 2005, 2012) explicitly argues for a genealogical connection between Lencan, Misumalpan, and Chibchan languages, within a putative “Lenmichí Microphylum”. In this genealogical unit, Lencan and Misumalpan languages would be more closely related to each other than either one would be related to Chibchan. Constenla Umaña (2005) proposes 85 Chibchan–Misumalpan cognate sets. Of these 85 sets, however, only six include both pan-Misumalpan and pan-Chibchan forms. In all other cases, the sets of alleged Chibchan–Misumalpan cognates only contain elements from either

- the Sumu-Ulwa and Cacaopera-Matagalpan branches of Misumalpan (but not Miskito) and Chibchan, or
- only one branch of Misumalpan and Chibchan, or
- all Misumalpan languages but only single, neighboring Chibchan languages.

This makes it difficult to exclude the possibility of borrowing in many cases. The strongest parallels that were identified by Constenla Umaña (2005) in favor of a genealogical Chibchan–Misumalpan connection are reproduced here in table 95 (Proto-Chibchan reconstructions are mine).

TABLE 95
PRESUMED MISUMALPAN–CHIBCHAN COGNATES PRESENTED BY CONSTENLA UMAÑA (2005)

Set	Proto-Misumalpan	Proto-Chibchan
1	*-ma, *man ‘thou’ (14)	*^{m}baʔ ‘you’
2	*li ‘liquid’ (20)	*^{n}diʔ ‘water, river’
3	*u ‘house’ (12)	*hu ‘house’
4	*nama ‘feline’ (12–3)	*^{n}da^{m}ba ‘feline$_{1}$’
5	*ai, *aima ‘maize’ (13)	*aiB ~ aBi ‘maize’
6	*ku ‘river turtle’ (19–20)	*k^{w}i ‘turtle’
7	Proto-Sumalpan *bu ‘two’ (15)	*^{m}buuʔ ‘two’
8	Proto-Sumalpan *usu ‘urine’ (16)	*hutsi ~ *huits ‘urine’

Misumalpan data from Constenla Umaña (2005), Proto-Chibchan reconstructions are mine. Page numbers in parentheses.

The forms ‘you’ and ‘water’ are indeed very stable parts of the vocabulary and regular sound correspondences attested in these terms are strong indicators of a genealogical relationship (Holman et al. 2008). Nevertheless, the scant systematic parallels between Proto-Chibchan and Proto-Misumalpan are probably not yet sufficient to argue a particularly close genealogical connection between the two language families. A problematic general issue in the case of Chibchan–Misumalpan coincidences is the difficulty to exclude an origin in language contact, given that the distribution areas are adjacent. While some contact may be relatively recent, the two language groups could also have influenced each other in a more distant past. For instance, Proto-Chibchan may have developed its largely trivocalic system under the influence of Misumalpan languages – trivocalic vowel inventories are infrequent and, in South America, are only attested in the western parts of the continent (cf. Adelaar 2012b: 603).

I do not aim to dismiss Constenla's 2005 proposal of a genealogical Misumalpan–Chibchan connection too hastily.[77] Notwithstanding, it seems that so far, evidence for a Misumalpan–Chibchan genealogical connection is not better than evidence for a genealogical connection between Misumalpan and Páez, an unclassified language from northwestern South America. Páez is spoken by some 40,000 people in the Cauca and Huila departments of southern Colombia (Adelaar & Muysken 2004: 618). Some (Proto-)Misumalpan and Páez sets, illustrating systematic correspondences between the two language groups are given in table 96.

TABLE 96
SOME (PROTO-)MISUMALPAN–PÁEZ LOOKALIKES

Set	(Proto-)Misumalpan	Páez	English
1	*tupal (CU 157), Proto-Sumalpan	*t*^{h}*ũ*^{ʔ}*wẽ*;	ear
2	*ubak (CU 155), Proto-Sumalpan	*uwe-*	to keep (Sp. *guardar*)/to hold
3	*paw (CU 158)	*beh*	red
4	*a (CU 158), Proto-Sumalpan	*ee*	blood
5	*wada 'hear' (CU 157)	*wẽsẽ*^{ʔ}*h-* 'listen'	to hear/listen
6	*walpa (CU 157)	*kwet*	stone
7	*lalalh (CU 152)	*lem-lem*, *lem*	yellow
8	*saŋ (CU 159)	*ʦẽj*	green
9	<sampapa> (Le 515)	*ʦme*^{h}*me*	butterfly
10	<sábaÿa> 'to pierce' (Le 536)	*swe*^{h}*de-* 'to bore'	to pierce/bore

Data from Lehmann (1920) (Miskito); Constenla Umaña 1987 (Proto-(Mi)sumalpan); Gerdel (2015) (Páez).

The vowel correspondence of Misumalpan *a and Páez *e* illustrated in table 96 seems to be quite recurrent, and it remains a matter of future investigation whether more evidence

[77] More relevant parallels might indeed eventually be discovered. Compare the correspondence of Proto-(Mi)sumalpan *t and Proto-Chibchan *k in the following forms (Proto-(Mi)sumalpan reconstructions by Constenla Umaña 1987): PM *tja 'nipple, teat' (158), PC *kaʔ 'breast, teat$_1$', PM *tu 'tongue' (155), PC *kuʔ(-Ba) 'tongue'; Proto-Sumalpan *ta 'mouth' (153), PC *kah 'open'; PS *tupal 'ear' (157), PC *kuh 'hear'; PS *titiŋ 'grandmother' (152), Constenla Umaña's (1981: 403) PC form *aˈkíkı 'old'. Page numbers in parentheses.

can be found to corroborate the existence of a possible genealogical Misumalpan–Paezan connection and whether this evidence remains more compelling than the evidence for a Misumalpan–Chibchan link. Some arguments in favor of the latter possibility are the following: (1) the recurrent Misumalpan–Páez *a*:*e* correspondences are found in sets that do not imply important semantic shifts (cf. Campbell 2003: 272; Adelaar 2013: 124); (2) the recurrent *a*:*e* vowel correspondence is also attested in stable parts of the vocabulary ('ear', 'to hear/listen', 'blood', 'stone', cf. Holman et al. 2008); (3) in contrast with the Misumalpan–Chibchan connection, recent borrowing can by and large be ruled out in the case of Páez and Misumalpan. This does of course not falsify the hypothesis of a genealogical Misumalpan–Chibchan connection, such as proposed by Greenberg (1987) and Constenla Umana (2005), but it suggests that the connection of Misumalpan with Páez might eventually turn out to be, in fact, closer or at least more easily demonstrable than a possible genealogical connection between Misumalpan and Chibchan.

3.2.2.2 Chibchan–Xinkan

Xinkan is an extinct, small language family formerly distributed in southeastern Guatemala (departments of Santa Rosa, most of Jutiapa, as well as the southern part of Jalapa) and adjacent areas of El Salvador and Honduras. It consists of the varieties Yupiltepeque, Jumaytepeque, Chiquimulilla, and Guazacapán (e.g., Campbell 1997: 166; Sachse 2010: 29). The main sources on Xinkan so far are Sachse's (2010) work on a Xinkan language described in the eighteenth century by the clergyman Manuel Maldonado de Matos (it is most similar to the Guazacapán variety), and Rogers' (2010) comparative grammar of Xinkan. Information about Xinkan toponyms is given by Campbell (1978). Evidence from Mayan loanwords seems to suggest that Xinkan-speaking groups dwell in Mesoamerica since before A.D. 500: Xinkan *k* in certain loanwords corresponds to Western Mayan *tʃ* or *tʃʔ*, and the sound change *k > *tʃ* or *tʃʔ* occurred before that date in Western Mayan (cf. Campbell 1972: 189; Kaufman & Norman 1984: 117–9; Sachse, forthc.). A relatively close connection between Chibchan and Xinkan has been proposed by Greenberg (1987: 106) and Voegelin and Voegelin (1977: 112). Among the few possible lexical parallels between the two language groups that I could identify is Xinkan *naki*, *nakʔi* 'chili pepper' (cf. Sachse 2010: 896) versus Proto-Chibchan *^{n}da^{ŋ}g 'salt', and Xinka *łuri* 'rabbit' (ibid.: 892), possibly referring to a kind of agouti, versus Chumulu (Dorasque) <*s*ole> 'red rabbit' (Sp. *conejo colorado*) (Pinart 1890: 10) and related forms in other Chibchan languages. None of these forms is exclusively shared by Xinkan and Chibchan, however: In the case of the 'salt' term, compare, for instance, Proto-Huave *ndɨ́kɪ 'sea' (Noyer 2012: 457) and Páez *nega* 'salt' (Gerdel 2015). For the 'rabbit' term (which may, in fact refer to a rodent like the agouti

or guinea pig, rather than to a rabbit), similar forms are also found in Lule <soly> ‘small rabbit’ (cf. Machoni 1732: 40) and in Kaingáng (Southern Jê) *for* ‘guinea pig (*Cavia* sp.)’ (cf. Jolkesky 2010: 252).

A South American origin of Xinkan-speaking groups has been suggested since the 1930’s at least (cf. Lothrop 1937: 202).[78] Indeed, in terms of lookalikes of Xinkan in non-Chibchan languages, the highest number of lexical parallels of Xinkan seem to exist in Proto-Southern Jê of southern Brazil. Some parallels are shown in table 97.

TABLE 97
SOME XINKAN–PROTO-SOUTHERN-JÊ LOOKALIKES

Set	18th-century Xinka	Proto-Southern Jê
1	*poːk* ‘pinetree, ocote’ (902)	*sɤg ‘araucaria (Araucaria angustifolia)’ (222); *sɔ ‘cedar (*Cedrela* sp.)’ (227)
2	*piːja* ‘leaf’ (901)	*sɛj ‘leaf’ (240)
3	***pur****i* ‘to burn’ (903)	*sũdsũɾ ‘blazing’ (222); *pũd ‘to burn (sg.)’ (254)
4	***paːma:*** ‘arm, wing’ (899)	*(jɔ-)**pɛ̃*** ‘arm, branch’ (225)
5	***ma****mi* ‘ear’ (893)	*bɛ̃ (-g) ‘to hear’ (249)
6	*mola* ‘moon, month’ (894)	*wɑɾ wɑɾ ‘moonlight’ (245)
7	*mi* ‘to say, s/he says’ (894)	*wĩ ‘to speak’ (238)
8	*kama* ‘blood’ (887)	*kɤwɛɟ ‘blood, bleed’ (256)
9	*tero* ‘to die’ (912)	*(kã̃g)teɾ ‘to die’ (247)
10	*kułku* ‘earthenware bowl, plate’ (889)	*kukɾũ{w} ‘pot’ (249)
11	*kˀunu* ‘cloud, shade, covered’ (890)	*kɾũj ‘fog’ (248)
12	*kˀɨsa* ‘bat’ (891)	*k{ɾ}ɤgsɛj ‘bat’ (247)
13	*ʔuj* ‘water’ (917)	*goj ‘water’ (220)
14	*ʔofo* ‘intestine’ (898)	*gɔɟ ‘intestine’ (243)
15	*ʃaha* ‘mouth’ (906)	*jã̃ ‘tooth, beak’ (224)
16	***ʃak****ˀajał* ‘thorn, spine’ (Sp. *espina*) (907)	***jɑg****sɯ ‘fishbone, spine’ (Port. *espinha*) (237)

[78] “It is believed probable that the Xinka of Guatemala were of South American extraction […].” (Lothrop 1937: 202)

17	***ʃaka**-ja* ‘to raise, lift, elevate’ (906)	*jɛ̃g ‘to stand, stand up (Port. *ficar em pé*)’ (239); *jɛ̃ ‘upright, erect, standing’ (236)

Sources are Sachse (2010) (Xinka) and Jolkesky (2010) (Proto-Southern Jê). Page numbers in parentheses.

Although a detailed investigation of these lookalikes and their possible origin is wanting, so far, the lexical coincidences between Xinkan and Proto-Southern Jê are more numerous than the parallels that I could identify between Xinkan and Chibchan (see above). A cursory look at Proto-(Macro-)Jê forms reconstructed by Davis (1968) and Nikulin (2015a) suggests that a comparison of Xinkan lexicon with Proto-Southern Jê yields more matching sets than a comparison of Xinkan and Proto-(Macro-)Jê. It is impossible so far to determine which exact scenario may account for this observation. A possible genealogical connection between Chibchan and Xinkan, however, is by no means falsified, as it were, by the Xinkan–Proto-Southern Jê lookalikes shown in table 97: In subsection 3.3 below, it will be argued that Chibchan has several cognate counterparts (lexical, but mostly grammatical elements) in (Macro-)Jê languages. Notwithstanding, if part of the Xinkan basic lexicon can indeed be linked with Southern Jê, and if Chibchan and Xinkan are visibly connected to Macro-Jê and Southern Jê, respectively, but not with each other, the nature of a possible connection between Chibchan and Xinkan still needs to be determined.[79] It goes without saying that Xinkan–Southern Jê parallels need further investigation before any strong claims can be made – in any case, however, they seem to be in line with a possible South American connection of Xinkan and are difficult to explain by recent language contact.

3.2.2.3 Chibchan–Lencan

Lencan is a small family of two extinct or nearly extinct languages, formerly spoken in eastern El Salvador and southern Honduras: Salvadorian Lenca (sometimes also referred to as Chilanga) and Honduran Lenca (e.g., King 2017: 5). The question of the external relations of Lencan has been considered to be “very important” (Stone 1948b: 205). Notwithstanding, it has remained somewhat open so far. Various authors have tried to connect Lencan – without providing convincing evidence, according to Constenla Umaña (2002: 189–90) – to different language groups, such as Mixe–Zoquean, Xinkan, Uto–Aztecan, and Mayan. Among the most frequently proposed, alleged relatives of Lencan are probably Chibchan and Misumalpan (e.g., Schmidt 1926: 206–8; Jijón y

[79] Unless the underlying phenomena are better understood, and provided that the parallels with Macro-Jê and Southern Jê parallels are not fortuitous, the relationship between Chibchan and Xinkan might tentatively be referred to as a ‘submerged relationship’.

Caamaño 1943: 11; Greenberg 1987: 382; all cited in Constenla Umaña 2002: 189). A connection of Lencan with Chibchan, and above all with Misumalpan, was also defended by Constenla Umaña himself (2002, 2012: 418). The strongest Chibchan–Lencan cognate sets proposed by Constenla Umaña (2005) are repeated here in table 98.

TABLE 98
PUTATIVE LENCAN–MISUMALPAN–CHIBCHAN COGNATES

Set	Salvadorian Lencan	Honduran Lenca	Proto-Chibchan
1	*lepa* ‘feline’	*lepa* ‘feline’	*^{n}da^{m}ba ‘feline$_1$’
2	***ma****nani* ‘you’	*amnan* ‘you’	*^{m}baʔ ‘you’
3	*mo* ‘smoke, steam’		*^{m}ba- ‘cloud, fog’
4	***nep****kuɾu* ‘nose’	***nep****sek* ‘nose’	*^{n}dii(k) ‘nose’
5	*pe* ‘two’	*pe* ‘two’	*^{m}buuʔ ‘two’
6	*pilil* ‘ashes, to burn’		*^{m}bũ^{n}d(ũ) ‘ashes’
7	*ʦʔahjaN* ‘green’		*tah(-ka) ‘raw’
8	*ʦʔana* ‘cold’		*ʦãĩh ‘cold’
9	*ʦʔu* ‘suck’; *ʦʔukiN* ‘teat’		*ʦuʔ ‘breast, teat$_2$’

Lencan data from Constenla Umaña (2005: 36–41); Proto-Chibchan reconstructions are mine.

The parallels shown in table 98 do involve some very stable vocabulary items, such as the terms for ‘nose’, ‘two’, and ‘you’. Nonetheless, it is difficult to exclude the possibility that Lencan lookalikes in Chibchan have different origins, for instance language contact, given the relatively close areas of distribution.

Comparing the inventory of Lencan lexicon with other Native American languages, several parallels can be identified with Taruma, a nearly extinct, unclassified language of the Guyana-Brazil border area (cf. Carlin 2007: 314–5). The homeland of the Taruma seems to have been close to the present-day city of Manaus (state of Amazonas, Brazil, near the confluence of the Negro and the Solimões River) (cf. Loukotka 1949: 53). Some sets of Lencan–Taruma lookalikes are given in table 99. The Taruma data were gathered, in 2015, by Sérgio Meira.

TABLE 99
SOME LENCAN–TARUMA LOOKALIKES

Set	Lencan	Taruma[80]
1	<guaz> ‘water’ (5), HL	*a ˈ**kwaʧu*** ‘creek, river’ (1)
2	<guag> ‘foot (Span. *pata*, *pie*)’ (11, 15), HL	*a ˈ**kwaʧa***, *a ˈ**kwaʧu**(ma)* ‘thigh’ (3)
3	<**gua**la> ‘hand’ (15), HL	*ˈa**kwa*** ‘arm’ (3)
4	<shug> ‘feather’ (11), HL	*u ˈ**ʤuk**a ‘leaf’* (*u* ‘tree’) (1)
5	<shali> ‘egg’ (11), HL	*a ˈkaɾa ˈ**ʤani*** ‘chicken egg’ (1)
6	<nepel> ‘tongue’ (14), HL; <nēbắļ> ‘tongue’ (712), SL	***nje ˈbe-**na* ‘tongue’ (3), cf. *njebe*, *ɲjebe* ‘lick’ (7)
7	<ŭa**šā́ga>** ‘foot’ (713), SL[81]	*u a ˈ**ʧaka*** ‘root’, cf. *u* ‘tree, wood’ (1)
8	*tik ‘face’ (105)	*a ˈ**ʧika**(ma)* ‘face’ (3)
9	<ála> ‘hair of the head’ (713), SL	*ˈada* ‘head’ (3)
10	<nĕp**kurū́**> ‘nose’ (712), SL; cf. Proto-Lencan *nep ‘nose’ (104)	*ˈasa **ku** ˈ**ru**tu* ‘nostril’, cf. *ku ˈrutu(ma)*, *kɨ ˈrɨtu(ma)* ‘hole’ (3)

Sources are Hernandez and Pinart (1897) (Honduran Lencan); Lehmann (1920) (Salvadorian Lencan); Arguedas Cortés (1988) (Proto-Lencan); and Meira (2015) (Taruma). Page numbers in parentheses.

A further potentially relevant parallel between Lencan and Taruma is attested in derivational morphology: a Taruma suffix *-na* is found, for instance, in *duri ˈki**na*** ‘entrance’, which is derived from *duri ˈki* ‘to close’. In Honduran Lencan, a suffix *-n* has been identified by King (2017: 30) in a form like *waki**n*** ‘grindstone’, from *wakik* ‘to grind’. Remnant morphology in the forms presented in table 99 needs to be determined in future studies. Little information is available so far on both Lencan and Taruma, and

[80] In Meira’s (2015) transcription, the following rules apply: “If a nasal vowel (with a tilde) is stressed, no stress mark is placed on it; the tilde itself has this value. If a vowel other than the nasal vowel is stressed, then it receives the acute accent and the nasal vowel the tilde, i.e. both diacritics co-occur.” (ibid.: 1). In the present version of Meira’s (2015) Taruma data, stress is indicated by <ˈ>. Also, Meira (ibid.) emphasizes that the Taruma wordlist in question “is preliminary, in the sense that the last Taruma speakers, who do not fully remember their language, may change their opinion on how to pronounce words and what they mean in future collaboration.”

[81] The element <ŭa> in Salvadorian Lenca <pī-**ŭa**šā́ga**>** ‘foot’ seems to recur in <**ŭa**pắla> ‘root’ (Lehmann 1920: 717) and may be cognate with Honduran Lencan <guag> ‘foot (Sp. *pata*)’ (Hernandez & Pinart 1897: 11, 15).

it goes without saying that the Lencan–Taruma lookalikes presented above do not allow for firm conclusions, except probably that so far, the evidence for a Lencan–Taruma connection does not seem less promising than the evidence for a Lencan–Chibchan connection (see table 98). This is suggested by (1) the fact that the lookalikes in table 99 are found in stable domains of the stable vocabulary such as 'hand', 'leaf', 'tongue', 'tooth' and 'water' (cf. Holman et al. 2008), and (2) by the distance between the distribution areas of Lencan and Taruma, which by and large excludes the possibility of recent borrowing.

3.2.3 South American languages

Many different authors have argued the existence of contact and genealogical relationships of Chibchan languages with South American ones, either in terms of contact (Uhle 1890; Jolkesky 2016) or in terms of genealogical relationships (e.g., Herzog 1886; Rivet 1924a; Greenberg 1987).

The present subsection briefly discusses some specific South American languages and their connections with Chibchan that have been proposed in previous publications: Páez, Chocoan, Pano–Tacanan, Tupían and Cariban.

3.2.3.1 Chibchan–Páez

Páez or Nasa Yuwe is a language spoken in southwestern Colombia, in the departments of Cauca and Huila (e.g., Adelaar & Muysken 2004: 618). It has been linked to Chibchan by different authors such as Rivet (1924a: 681) and Jijón y Caamano (1943: 418). Greenberg (1987: 106) classified it within an alleged Chibchan–Paezan subgroup of 'Amerind'. Chibchan and Páez have rather different phonological systems, and, what is more important here, systematic correspondences are wanting. Notwithstanding, a certain (probably contact) connection between both language groups may have existed; it is reflected in cases such as the Páez diminutive marker *-kwe* (cf. Rojas Curieux 1998: 121) versus Proto-Chibchan *k^{w}a 'seed, fruit', and Páez *nega* 'salt' versus Proto-Chibchan *^{n}da^{ŋ}g 'salt'.

In terms of external connections of Páez with non-Chibchan languages, Jolkesky (2017) proposes the existence of a genealogical relationship of Páez with Oto–Manguean languages. In subsection 3.2.2.1, some preliminary evidence for recurrent Páez–Misumalpan vowel correspondences has been provided, which seems to be reflected in several basic lexical items. This evidence will not be repeated here. A genealogical connection of Páez may also exist with Andakí. This language was formerly spoken in the area of the upper Magdalena river, and in the first half of the twentieth century, there may still have been some Andakí in the headwater areas of the

Fragua and Mandiyaco rivers in southern Colombia (Friede 1953: 44–65). The Andakí were famous for their fierce resistance and guerilla struggle against invaders of their area (ibid.: 98–107). All available Andakí forms so far are contained in Coronas Urzúa (1995).

Andakí was classified by Greenberg (1987: 106) as nuclear Paezan within the Paezan group of Chibchan–Paezan. Rivet (1924b), Adelaar and Muysken (2004: 140), and Jolkesky (2016: 539–40) mention several Páez–Andakí lookalikes. Indeed, coinciding forms can be found in the stable parts of the lexicon, suggesting the possibility of a genealogical relationship between the two languages. Table 100 repeats three parallels identified by Adelaar and Muysken (2004) (the forms for 'tongue', 'ear', and 'to sleep'), three of Jolkesky's (2016) additional findings (the terms for 'face', 'tail', and 'stone'), and gives five further Páez–Andakí parallels. The Andakí data in table 100 are from Rivet's (1924b) work and are ultimately based on Albis (1860/61). Rivet identifies morpheme boundaries in these Andakí materials which are kept as such in table 100.

TABLE 100
SOME PÁEZ–ANDAKÍ LOOKALIKES

Set	Páez	Andakí
1	*kuse* 'hand'	<gosoa> 'arm' (Rib 101)
2	*kwet* 'stone'	<guatihi> 'stone' (GT)
3	*t^hũ^{ʔ}wẽ*, *t^hũ^{ʔ}wã* 'ear'	<sungua-xo> 'ear' (Rib 101), <chunguahe> (GT)
4	*t^hune* 'tongue'	<sonae> 'tongue' (Rib 108), <shonae> (GT)
5	*^{n}de^h* 'to sleep'	*da-*, *bo**nda***- 'to sleep' (AdM 140)
6	*mez* 'tail'; *e^{ʔ}s-* 'after, behind'	<ma-esegua> 'tail' (Rib 101), <maszengua> (GT)
7	*knene* 'forehead'	<kina-xi> 'head' (Rib 104), <quinaxi> 'head', <quiexa> 'hair' (GT)
8	*d^{j}i^{ʔ}p* 'face'	<chipina> 'face' (GT)
9	*d^{j}i^{ʔ}t^h* 'bone'	<chicazo> 'bone' (GT)
10	***ɸi**w* 'seed', *ja**ɸ*** 'eye', *ka**ɸ*** 'hole'	<si-**fi**> 'eye' (Rib 100), <szi**fi**> 'eyes', <chimbu**fi**> 'navel' (GT)

11	***ĩk**^{h} wala* ‘lake’	<jixe, jexe> ‘water’ (GT)

Sources are Rivet (1924b) and Gómez and Torres (2012/13) (Andakí), and Gerdel (2015) (Páez). Page numbers in parentheses.

A detailed assessment of remnant morphology remains forthcoming. In the parallels shown in table 100, Andakí <a> regularly corresponds to Páez *e* in sets 1 to 7. This systematic vowel correspondence in basic vocabulary items suggests that the relatively close genealogical relationship between Páez and Andakí, proposed by Greenberg (1987: 106) might indeed turn out to be valid. A similar vowel correspondence was also identified above, in subsection 3.2.2.1, where it was shown that Misumalpan *a* likewise corresponds to Páez *e*. It may thus be argued that in the sets in question, Páez *e* is innovative, whereas Misumalpan and Andakí *a* are retentions. So far, I could find more potential evidence for a genealogical connection between Páez and Andakí (and Misumalpan) than for a genealogical connection between Páez and Chibchan. It goes without saying that the exact relationships between Páez, Andakí, and Misumalpan need to be determined in future investigation.

3.2.3.2 Chibchan–Chocoan

The distribution of Chocoan languages is somewhat interspersed, at present, between the Chibchan areas of Central America and Colombia. Chocoan languages are spoken from eastern Panama to western Colombia and northwestern Ecuador. This language family consists of two languages, Waunana (or Wounaan meu) and Emberá. The latter has some 60,000 speakers and forms a continuum of different dialects mainly spoken in the Colombian departments of Chocó, but also in Antioquia, Risaralda, Córdoba, Cauca and Nariño (Aguirre Licht 2006; Hoyos Benítez 2000: 73). In Ecuador, the Río Verde dialect of Emberá is spoken by some 1,000 people (Hoyos Benítez 2000: 73). Waunana has some 8,000 speakers living in the Colombian department of Chocó (Mejía Fonnegra 2000: 85), plus some in eastern Panama. Extinct Cueva from eastern Panama was probably also a Chocoan language (Loewen 1954: 4–5).

Several authors have more or less tentatively proposed the existence of a relatively remote genealogical link between Chibchan and Chocoan languages (e.g., Lehmann 1920: 57; Constenla Umaña & Margery Peña 1991: 172). Greenberg (1987: 106) classified Chocoan as Paezan within his Chibchan–Paezan group. Table 101 reproduces the strongest cases among the fifteen putative cognate sets that were very

carefully and tentatively proposed by Constenla Umaña and Margery Peña (1991), in favor of a possible genealogical connection of Chibchan with Chocoan.[82]

TABLE 101
SOME POSSIBLE CHIBCHAN–CHOCOAN COGNATES.

Set	Proto-Chocoan	Proto-Chibchan
1	*da first person plural	*^{n}daH(^{n}d) 'I'
2	*ˈbɯ second person singular	*^{m}baʔ 'you'
3	*i third person	*hi ~ *iʔ 'that'
4	*kãi 'sleep'	*kap- 'to sleep'
5	*be 'maize'	*aiB ~ aBi 'maize'
6	***khuri**ˈhia Proto-Emberá 'paca, agouti'	*ku^{n}di ~ *kui^{n}d 'rodent, large'

These possible cognates were presented as such by Constenla Umaña and Margery Peña (1991: 172). Proto-Chibchan reconstructions are mine.

The parallels shown in table 101 are probably not sufficient to substantiate a close genealogical Chibchan–Chocoan connection. The existence of a close genealogical connection between Chibchan and Chocoan was refuted by Holmer (1947: 206). However, there are some parallels of Chocoan and Chibchan languages that indicate contact. These parallels are shown in table 102.

TABLE 102
SOME CHIBCHAN–CHOCOAN PARALLELS, PRESUMABLY REFLECTING CONTACT

Chocoan	Chibchan
Proto-Chocoan *de 'house'	Bribri [tɛ] 'house', used in shamanic chants (this coincidence has been mentioned, for instance, by Constenla Umaña & Margery Peña 1991: 139).

[82] The authors state: "No es el objetivo de este artículo el tratar las relaciones de la familia chocó con otras agrupaciones de lenguas, pero en el transcurso de ella se han notado algunas semejanzas con la familia chibcha que, para nosotros, han hecho más sugestiva y digna de explorarse la hipótesis de las relaciones entre ambas familias que los elementos de juicio hasta el momento por los principales proponents de la misma (Lehmann 1920, Greenberg 1987)." (Constenla Umaña & Margery Peña 1991: 172)

Proto-Emberá *ˈkẽ ‘vulva’	Bribri *kĩ*, Cabécar *kɪ* ‘vulva’ (Constenla Umaña & Margery Peña 1991: 172)
Cueva <haba> ‘basket’ (Romoli 1987: 73), itself probably borrowed from Kogi /gaˈma/ a kind of carrying bag (Span. *mochila*) (Ortíz Ricaurte 2000: 765) or a related form	San José Cabécar [hába] ‘basket’ (Margery Peña 1989: 397), a loanword from Chocoan
Reduplication plays an important role in Epena Pedee (Saija), where it indicates a higher or lower degree of the quality referred to by an adjective or color term (Harms 1994: 39–40)	In Cabécar and Teribe, reduplication in adjectives or color terms may either express a higher or a lower degree of intensity of the quality referred to (Margery Peña 1989: lii–liii; Quesada 2000a: 88–9)
Panamanian Emberá *ʧaˈwala* ‘harpoon’	Kuna *esawala* ‘iron point or harpoon (fish spear) with an iron point; heron’s beak’ (cf. Urban, forthc.a)
Northern Emberá *kamisu* ‘paddle; kitchen implement’	Kuna *kammi* ‘paddle’ (cf. Urban, forthc.a)

The evidence shown in table 102 illustrates that Kuna, which is still spoken in the vicinity of Chocoan languages, but also Central American Chibchan languages such as Bribri, Cabécar, and Teribe/Térraba have been in contact with Chocoan languages in the past. In the case of the ‘basket’ term in table 102, a borrowing chain from Chibchan (Kogi or Proto-Chibchan) into Chocoan (Cueva or another Chocoan language) and back into Chibchan (San José Cabécar) might be proposed, given that this is the only possibility to account for the correspondence of Kogi *g* in *gama* ‘*mochil*a’, Cueva <h> in <haba> ‘basket’, and San José Cabécar *h* in [hába] ‘basket’ – the regular reflex of a putative Proto-Chibchan form *^{ŋ}ga^{m}ba ‘basket’ in Kogi is *gama*, and it would have yielded *ʤaba** or a similar form in Cabécar, not *hába*. Instead, a sound change *g > *h* is attested in Chocoan, and the Cabécar form is therefore most probably a loan from a Chocoan language (cf. Pache 2016b: 116). The parallels between Chocoan and certain Central American Chibchan languages are somewhat unexpected, given the fact that at present, Chocoan languages are mostly spoken in western Colombia, far away from western Panama and eastern Costa Rica.

Among the most striking parallels between Chibchan and Chocoan are probably those involving relational/possessive markers; these parallels are shown in table 103:

TABLE 103
POSSESSIVE/RELATIONAL ELEMENTS SHARED BY CHOCOAN AND CHIBCHAN LANGUAGES

Chocoan	Chibchan
Epena Pedee *ne-*, attached to nouns and verbs; with nouns, this element allows it to omit a possessor which is unknown or irrelevant, as in *ne-ˈʔɨmɨ* 'egg' and *ne-ˈp^{h}ono* 'flower' (Harms 1994: 41–2). Generic *ne-* in Epena Pedee can only be combined with a limited set of nouns; the form *ne-ˈk^{h}iru** 'a leaf', for instance, is ungrammatical (Harms 1994: 42).	Proto-Chibchan relational$_1$ *^{n}d(i)-, as reflected in: Damana *ni-* in ***ni**kuma* 'egg' (Huber & Reed 1992: 260) (compare Boruca *kúp* 'egg, testicle' in Quesada Pacheco & Rojas Chaves 1999: 153, 193); Kogi *ne-* in /ˈneuma/ 'brother-in-law' (Ortiz Ricaurte 2000: 761) (compare Proto-Chibchan *^{n}d-u^{m}ba ~ *u^{m}ba 'brother-in-law'); Boruca *ɾ-* in ***ɾ**unkáx* versus *unkáʔ*, both 'father-in-law' in this language (Quesada Pacheco & Rojas Chaves 1999: 191).
Waunana *-c(i)* can be used as a possessive marker with personal pronouns, for instance in *mə-**ci** eum(u)* or *mə eum(u)* 'my younger brother' or in *i-**c** nemek* 'his/her food' (Holmer 1963: 113).[83]	Proto-Chibchan relational$_2$ *ʦ(i)-, as reflected in: Cabécar *-s* in *huɽa**s** kʔǒ* 'finger', literally 'leaf (*kʔǒ*) of the hand (*huɽa*)' (Margery Peña 1989: 419) Muisca <-s> in <cha-**s** gue> 'the man's (<cha-s>) house (<gue>)' González de Pérez 1987: 73, 137, cited in Adelaar & Muysken 2004: 102) Rama *si* in *up**si**-ri* 'tears' literally 'water (*ri*) of the eye (*up*)' (cf. Craig 1989: 66).

In table 103, it is remarkable that reflexes of *^{n}d(i)- and related forms are attested in single, non-contiguous forms in Chocoan. In Chibchan, a similar phenomenon can be

[83] <c> in Holmer's (1963) transcription corresponds to *ɛ* in Mejía Fonnegra's (2000) analysis.

observed (compare, for instance, Damana ***ni****kuma* 'egg' (Huber & Reed 1992: 260); for more examples see subsection 2.2). Reflexes of *ʦ(i)- and related forms are only found in contiguous constructions both in Chocoan and in Chibchan (for instance, Cabécar *huɻa****s*** *kˀó* 'finger', literally 'leaf (*kˀó*) of the hand (*huɻa*)' (Margery Peña 1989: 419, Margery Peña 1989: 419)). The origin of these parallels needs further investigation.

Notwithstanding these Chibchan–Chocoan coincidences, an alternative, relatively close genealogical connection of Chocoan has been argued to exist with Pumé or Yaruro, a language of the lower Apure, Arauco and Cinaruco rivers and the Orinoco, in the Apure state of Venezuela (Pache 2016b). The fact that the relationship between Chocoan and Pumé is not fortuitous is suggested by the recurrent character of sound correspondences reflected in shared (near-)homophones. Table 104 shows two shared near-homophones in Pumé and Epena Pedee (or Saija, a southern Emberá dialect, Chocoan) (for the explanation of (fossilized) morphological elements in these and other Pumé–Chocoan sets, see Pache 2016b).

TABLE 104
NEAR-HOMOPHONES SHARED BY CHOCOAN AND PUMÉ

Set	Chocoan (Epena Pedee)	Pumé
1	***hẽra*** *bee-* 'to lie down'	*ãrẽ* 'to lie down'
2	*ẽra-* 'to untie'	*hare-* 'to untie'
3	*ɨa-* 'to keep, retain'	*gõa-* 'to have, own, possess, hold, carry-in-hand'
4	*hɨã-* 'to shut, close'	***goa****di-* 'to hide, conceal'

Sources are Quiro Dura and Harms (2015) (Epena Pedee) and Dyck and Dyck (2015) (Pumé).

The existence of a genealogical link between Pumé and Chocoan is suggested by the fact that systematic sound correspondences are reflected in basic vocabulary items such as body-part terms, shown in table 105.

TABLE 105
PUMÉ AND CHOCOAN BODY-PART TERMS

Set	Chocoan (Epena Pedee)	Pumé
5	*ˈtau* ‘eye’	*da* ‘eye’ (used in complex forms)
6	*tautʰu* ‘forehead’	*dacc͡o* ‘eye, face’
7	*tɾu* ‘tail’, Waunana *-duɾi* ‘tail’ (204)	*duɾi* ‘after’
8	*ˈwaa* ‘blood’	*goe* ‘blood’, *gõã* ‘meat, flesh’
9	*e* ‘skin’	*i* ‘skin’
10	*kẽˈbɨ* ‘nose’ (northern Emberá)	*ĩbu* ‘nose’
11	*iˈsia* ‘wing’	*ic͡çi* ‘hand’

Sources are Holmer (1963) (Waunana), Quiro Dura and Harms (2015) (Epena Pedee), and Dyck and Dyck (2015) (Pumé). Page number in parentheses.

The fact that recurrent sound correspondences are attested in the terms for ‘eye’, ‘nose’, ‘blood’, ‘skin’, ‘hand’ (table 105) and ‘water’ (cf. table 106 below) deserves particular attention, since across different languages of the world, these elements have proven to be among the most stable parts of the vocabulary (cf. Holman et al. 2008). Pumé and Chocoan also share some morphological parallels, for instance in the context of postposed elements indicating gender (Pache 2016b: 120–2). Pumé–Chocoan sound correspondences have been discussed in detail in Pache (2016b) and are not repeated here. Instead, I will discuss some particular sound changes in what follows: The regular sound correspondences attested between Pumé and Chocoan do not only suggest the existence of a genealogical connection between both language groups, but they also allow it to reconstruct some elements of the shared ancestor language (see subsection 3.1.1.1 above). For instance, the forms shown in table 106 suggest that the putative ancestor language, Proto-Pumé–Chocoan, had a phoneme *k, the outcome of which is zero in Pumé and *k* in Chocoan languages (that is, in table 106, Chocoan *k* regularly corresponds with Pumé zero).

TABLE 106
SEVEN SETS ILLUSTRATING THE DEVELOPMENT OF PROTO-PUMÉ–CHOCOAN *k

Set	Chocoan (Northern Emberá)	Chocoan (Waunana)	Pumé
12	*ka-* 'to weave'		*a-* 'to weave'
13	*koˈɾe* 'crocodile, alligator'	*nem-**kõre*** 'snake' (223)	*aɾi* 'alligator'
14	*kiˈma* 'spouse'		*ɛba-* 'marry'
(10)	*kẽˈbɨ* 'nose'	*kẽwa* 'nose' (215)	*ĩbu* 'nose'
15	*kue* 'rain'		*ui* 'water'
16		***ko**dama* 'cousin, relative' (216)	*o* 'relatives, kinsmen'
17		***kõĩ**ra* 'married man' (216)	*ɔĩ̃* 'man'

Sources are Pardo Rojas (2015) (Northern Emberá), Holmer (1963) (Waunana), and Dyck and Dyck (2015) (Pumé). Page numbers in parentheses.

Thus, Pumé is apparently innovative in the sense that the reflex of Proto-Pumé–Chocoan *k is zero in this language, whereas it remained as such in Chocoan.

The sets in table 107 instead suggest that Chocoan *ko* or *ku* regularly corresponds with Pumé *ko*. Given this recurrent correspondence, it is postulated that *ko* or *ku* in Pumé and Chocoan (*go* in intervocalic position in Northern Emberá) reflect a Proto-Pumé–Chocoan labialized velar *k^{w}.

TABLE 107
THREE SETS ILLUSTRATING THE DEVELOPMENT OF PROTO-PUMÉ–CHOCOAN*k^{w}

Set	Chocoan (Northern Emberá)	Pumé
18	*kõ-* 'to cut'	*koa-* 'to cut'
19	***kua**ˈra* 'yellow'	***koɛ**-koɛ-a* 'red'[84]
20	*õˈ**goa*** 'raw'	***koɛ**-ʧʰa* 'raw'

Sources are Pardo Rojas (2015) (Northern Emberá) and Dyck and Dyck (2015) (Pumé).

[84] Compare reduplication in Emberá color terms discussed in table 102 above.

Finally, *go* or *gõ* in Pumé corresponds with *w* or *u* in Chocoan languages, as presented in table 108. This may reflect either Proto-Pumé–Chocoan *ŋg^{w} or *w.

TABLE 108
THREE SETS ILLUSTRATING THE DEVELOPMENT OF PROTO-PUMÉ–CHOCOAN *g^{w}

Set	Chocoan (Epena Pedee)	Chocoan (Waunana)	Pumé
21	*ua-* 'carry-in-hand'	*wai-* 'to have, possess' (237), *war-* 'to carry, take' (238)	*gõa-* 'to have, own, possess, hold, carry-in-hand'
(8)	*ˈwaa* 'blood'		*goe* 'blood', *gõã* 'meat, flesh'
22	*wãkʰa-* 'to cover'		*goadi-* 'to hide, conceal'

Sources are Quiro Dura and LHarms (2015) (Epena Pedee), Holmer (1963) (Waunana), and Dyck and Dyck (2015) (Pumé). Page numbers in parentheses.

The regular correspondences between Chocoan and Pumé, and developments discussed above corroborate the systematic connection between both language groups. Whereas speakers of Chibchan and Chocoan languages have been living in close contact, recent, intense language contact as the origin of the Pumé–Chocoan parallels given in tables 104 to 108 can be ruled out, given that the distribution areas of Pumé and Chocoan groups are divided by two to three Andean cordilleras: whereas Chocoan languages are spoken in western Colombia and adjacent areas of Panama and Ecuador, Pumé is spoken in the western savannah plains of Venezuela between the Apure, Arauca, Cinaruco, and Orinoco. In sum, the arguments for a Chibchan–Pumé connection seem to be more compelling than those in favor of a Chocoan–Pumé connection for the moment.

3.2.3.3 Chibchan–Pano-Tacanan

Several researchers have noticed the existence of lexical parallels between Chibchan languages on the one hand, and Panoan and/or Tacanan languages on the other hand (e.g., Uhle 1890: 467; Holt 1976, 1986: 229–260). Pano–Tacanan languages are spoken in a large area covering parts of Peru, Brazil, and Bolivia. The distribution of Tacanan languages is mainly attested within the boundaries of Bolivia at present and thus somewhat more to the south than the distribution of Panoan languages (cf. Adelaar & Muysken 2004: 418; Adelaar 2012b: 585). There are some thirty Panoan languages spoken in the Peruvian departments of Loreto, Ucayali, Huánuco, Madre de Dios, and Cuzco, in the Brazilian states of Acre, Amazonas, and Rondônia, and in the northern Bolivian departments of Beni and Pando (Valenzuela & Guillaume 2017: 3). The five Tacanan languages that are still spoken are presently distributed in northern Bolivia and, in the case of Ese Eja, also in adjacent areas of Peru (ibid.: 4). The genealogical relationship between Panoan and Tacanan was probably first explicitly proposed by Schuller and Gusinde (1933: 480) and has been discussed since then (see, e.g., Guillaume 2008: 7). Adelaar and Muysken (2004: 419) and Valenzuela and Zariquiey (2015) are among those authors who consider Pano–Tacanan to be a valid genealogical unit.

The first author to observe lexical parallels between Chibchan and Panoan was Uhle (1890: 467) who presented the following Muisca–Kulina lexical correspondences (table 109).

TABLE 109
MUISCA–KULINA (PANOAN) LOOKALIKES

Muisca	Kulina	English
<gua>	<ghüma>	‘fish’
<chue>	<tschuma>	‘breast’
<sie>	<oschý>	‘moon’
<yba>	<ymý>	‘blood’
<tschay>	<sue>	‘bird’

Source: Uhle (1890: 467).

Uhle (1890: 468) remarks that Kulina <ghüma> ‘fish’ relates to Muisca <gua> ‘fish’ in a similar way that Kulina <tschuma> ‘breast’ relates to Muisca <chue> ‘breast’, but

adds that these correspondences are not very important ("aber das ist keine Gleichung tieferer Wichtigkeit"), and that there are no more compelling similarities ("beweisendere Aehnlichkeiten sind nicht vorhanden"). A possible relationship between Panoan and Chibchan languages was taken up by Swadesh (1954: 325) who compared two words of Kashinawa and Muisca, namely Kashinawa <mɨʃi> 'black' and <kiʃi> 'foot' with Muisca <mɨhsio> 'black' and <kihca> 'foot'. Holt (1986: 229–260) argued for a possible genealogical connection between Chibchan and Pano–Tacanan languages.[85] He presented 72 supposed sets of Proto-Chibchan cognates in Proto-Panoan and/or Proto-Tacanan, consulting 504 Tacanan reconstructions of Girard (1971) and 512 Panoan reconstructions of Shell (1965) (cf. Holt 1986: 229–58). Holt identified parallels with Proto-Tacanan in all 72 cases, but only in 16 cases for Proto-Panoan. Thus, 56 items are exclusively shared between Proto-Chibchan and Proto-Tacanan.

Indeed, Chibchan–Tacanan parallels and the possibility of a genealogical relationship between the two language groups were commented upon by Constenla Umaña (1981: 14) in a relatively favorable way.[86] The five Chibchan–Tacanan parallels that Constenla Umaña (ibid.) provides are reproduced in table 110 (I only give evidence from Kogi and Tacana here).

[85] It is striking that additionally, both Holt (1986, 1997/98) and Swadesh (1954: 324–5) see a genealogical relationship of Chibchan with Uto–Aztecan, and that Wistrand-Robinson (1991) argues a genealogical connection between Panoan and Uto–Aztecan. In terms of triangulation, Wistrand-Robinson's apparently independent proposal seems to corroborate an alleged genealogical connection of Chibchan with Pano–Tacanan languages: If a genealogical connection exists between Chibchan and Pano–Tacanan, and between Chibchan and Uto–Aztecan, a genealogical connection should likewise be identifiable between Uto–Aztecan and Pano–Tacanan. The putative close genealogical relationship of Chibchan with Uto–Aztecan languages, however, has been rejected by Constenla Umaña (2012: 418). Notwithstanding, some of the Chibchan–Uto-Aztecan lexical parallels identified by Holt (1986, 1997/98) should probably not be too hastily dismissed, for instance those involving Proto-Uto–Aztecan *mo 'cloud' or *nɨʔ 'I'. It is unclear how these and other coincidences with Proto-Chibchan can be accounted for in the context of the present study.

[86] "Similarly, the Tacanan languages […] present lexical resemblances to the Chibchan languages that seem to me at least as promising, or perhaps more so, than those that I have been able to detect in many proposed members of Macro-Chibchan." (Constenla Umaña 1981: 14)

TABLE 110
KOGI–TACANA LOOKALIKES

Kogi	Tacana	English
<muli>	<moro>	'ashes'
<abi>	<ami>	'blood'
<maui>	<bo>	'cloud, fog'
<tai>	<te>	'planted field'
<mai>	<mia>	'thou'

Source: Constenla Umaña (1981: 14).

Constenla seems to have kept his positive opinion towards a Chibchan–Tacanan connection for a relatively long time, since later (Constenla Umaña 2005: 9) he still considers Holt's (1986) proposal of a genealogical link with Pano–Tacanan "interesting".[87] Apart from the parallels shown in table 110, I did not find sufficient indications, however, for a close genealogical connection between Chibchan and Tacanan, through comparing my own reconstructions with Girard's (1971) Proto-Tacanan forms. Among the additional parallels (which in most cases are not exclusively found in Proto-Tacanan) are Proto-Chibchan *ku(^{n}d) 'feline$_2$' versus Proto-Tacanan <*kuři> 'jaguar, feline' (Girard (1971: 90); a similar parallel has already been identified by (Holt 1986: 255)),[88] Proto-Chibchan *pu^{n}d- 'dry$_2$' versus Proto-Tacanan <*muři> 'dry' (Girard 1971: 102); Proto-Chibchan *^{m}ba^{n}d- 'tasty' versus Proto-Tacanan *banu 'salt' (Girard (1971: 61); a similar parallel has already been identified by Holt (1986: 252)), and Proto-Chibchan *^{m}bu^{n}d-/*^{m}bus- 'woman' versus Proto-Tacanan *puna 'woman' (Girard (1971: 113); a similar parallel has already been identified by Holt (1986: 255)).

If the similarities between Chibchan and Pano–Tacanan are non-fortuitous, the fact that Holt (1986: 229–60) identified many more parallels with Proto-Tacanan than with Proto-Panoan may be explained by different scenarios. In one scenario, Proto-Chibchan was in contact with Proto-Pano–Tacanan, and Panoan lexicon underwent heavy changes after the separation of Proto-Panoan and Proto-Tacanan. The alternative possibility would be that Proto-Chibchan was in contact with (Proto-)Tacanan after a

[87] "Por el momento, sin embargo, ninguno de los múltiples parentescos [...] se ha probado (si bien en algún caso, particularmente en el de la propuesta de Holt sobre la relación con el pano-tacana, se han aportado indicios interesantes) [...]" (Constenla Umaña 2005: 9).
[88] The Proto-Tacanan symbol <*ř> may refer to a "flap articulation simultaneous with friction made by the tongue blade" (Girard 1971: 34).

split of Proto-Pano–Tacanan into Proto-Panoan and Proto-Tacanan. The latter scenario is remarkable in that the distribution area of Tacanan languages is further to the south than the Panoan area, and thus geographically more distant from the present-day Chibchan area. Also, the latter scenario would imply that the ancestor language of Proto-Chibchan was still spoken in South America at a relatively recent period, namely after Proto-Pano–Tacanan split into Proto-Panoan and Proto-Tacanan.

3.2.3.4 Chibchan–Tupían–Cariban

A genealogical connection between Chibchan and languages of eastern South America has been suggested from the beginning of comparative Chibchan studies, namely by Herzog (1886) who argued for a close genealogical connection between Central American Chibchan languages such as Bribri, Térraba, and Guatuso and two language groups of eastern South America: (1) Tupí–Guaraní, including Tupí, Tupinamba, Guaraní, Wayampi, and Omagua,[89] and (2) Cariban, including, among other languages, Macushí, Kari'na, but also Warao.[90] The only predecessor in this matter explicitly mentioned in Herzog's (1886) paper was Brinton (1883: xiv–xv, cited in Herzog 1886: 625) who had stated before him that "the great Tupi-Guarani stem, [...] has extensive affiliations in Central America". Also, although this is not explicitly stated by Herzog (1886), it should likewise be noted here that the Jesuit scholar Hervás y Panduro had ventured a genealogical connection of Guaymí and the language of Darien (Kuna?) with Island Carib, a Maipuran language highly influenced by Cariban (Hervás y Panduro 1784: 72; 1800: 281). Whereas Herzog (1886: 625) interprets the connection of Central American Chibchan languages with Tupí–Guaraní and Cariban as genealogical, he also sees language contact at the origin both of parallels between Central American Chibchan languages and Ulwa, Miskito (both Misumalpan), and between the Chibchan languages in question and Chocoan languages.

Tupí–Guaraní is a family of the Tupían stock (Rodrigues & Cabral 2012: 498). Whereas a homeland of Proto-Tupían has been argued to have been in an area corresponding to the present-day Brazilian state of Rondônia (ibid.: 499–500), Tupí–Guaraní languages widely expanded across South America, above all in relatively recent times. As a result, Tupí–Guaraní languages are still spoken in an area extending from the upper Amazon in the northwest to present-day Paraguay in the south and the Guyanas in the northeast (cf. Rodrigues & Cabral 2012: 495–501). Herzog (1886: 626)

[89] For a discussion of the origin of Omagua, compare Michael (2014).

[90] At present, Warao is considered an unclassified language. Literally, Herzog (1886: 625) states: "Nicht durch Zufall, nicht durch späteres Eindringen dürften jedoch die häufigen und zahlreichen Coincidenzen der costaricensischen Dialecte mit dem Tupi Brasiliens und den caribischen Dialecten Süd-Amerikas zu erklären sein. […] Ich wage es nun zu behaupten, 'die Bribri-Dialecte Costa-Ricas sind mit dem Tupi und dem Caribischen nahe verwandt'."

observes that certain toponyms from Costa Rica have counterparts in Tupí, for instance, *Arariba* and *Cururu* in Costa Rica, which he compares to the toponyms *Araripe* and *Cururú* in Brazil (he probably found the latter toponyms in Martius (1867: 491, 498)). Besides that, the evidence given by Herzog (1886) are 40 sets of Central American Chibchan–Tupí-Guaraní lookalikes. He sees them as proofs ("Belege") of a genealogical relationship (Herzog 1886: 625). Some examples of Herzog's evidence are cited in table 111.

TABLE 111
SOME PUTATIVE CENTRAL AMERICAN CHIBCHAN–TUPÍ-GUARANÍ COGNATES.

Tupí-Guaraní	Chibchan
<bora> a kind of bee (Tupí)	<bur, bor> 'bee' (Bribri)
<yvy> 'sand', <uvu> 'earth' (Tupí)	<up> 'sand' (Boruca)
<ku, cu> 'tongue' (Guarani)	<cu, ku> 'tongue' (Bribri)
<ua> 'fruit' (Tupí)	<va> 'fruit' (Bribri)
<quibira, kebyra> 'brother' (Tupí)	<s'quibire> 'brother' (Chirr.)[91]
<curú> 'pitcher, pot' (Ger. *Krug*, *Topf*) (Tupí)	<curéh> 'pitcher, pot' (*Krug*, *Topf*) (Boruca)
<tin, ting> 'nose' (Tupí)	<táin> 'nose' (Guatuso)

Cognates presented as such by Herzog (1886: 625–6).

Comparing the two reconstructed ancestor languages, Proto-Chibchan and Proto-Tupían, a few lookalikes can indeed be identified. These cases are shown in table 112. As such, they are not sufficient to postulate a close genealogical link between Chibchan and Tupían languages, however.

[91] This abbreviation used by Herzog (1886) may refer to the Cabécar variety of Chirripó.

TABLE 112
SOME PROTO-CHIBCHAN–PROTO-TUPÍAN LOOKALIKES

Set	Proto-Tupían	Proto-Chibchan
1	*ka allative postposition (RC 517)	*ka allative/dative
2	*ʔek^{w} ‘tuber, edible root’ (Rob 182)	*i(k) ‘manioc’
3	*kʲũ ‘tongue’ (Si 229)	*kuʔ(-Ba) ‘tongue’
4	*ŋam ‘breast’ (RC 507)	*kãʔ ‘breast, teat$_1$’
5	*ŋup ‘louse’ (RC 507)	*kũʔ ‘louse’
6	*pap ‘to die’ (RC 503)	*kap- ‘to sleep’
7	*pe punctual locative, dative (RC 517)	*ki locative$_2$
8	*atsuk ‘to bathe’ (Rob 178)	*suh(k) ‘to wash’
9	*tsup ‘to see’ (RC 504)	*su ‘to see’

Proto-Tupían reconstructions from Silva (2007); Rodrigues (2007); and Rodrigues and Cabral (2012); Proto-Chibchan reconstructions are mine.[92]

Correspondences between Chibchan *k and Tupían *p as in *kap- ‘to sleep’ versus *pap ‘to die’ are not necessarily problematic – a similar variation between bilabial and velar stops has also been postulated for Proto-Chibchan (see above, subsection 2.4.3.1), and a similar variation may also have existed in Proto-Tupían: compare Proto-Tupían *kɨt ‘green, unripe’ (Rodrigues & Cabral 2012: 503) versus *pɨt ‘raw’ (Rodrigues 2007: 174).[93] Another formal parallel in certain (Proto-)Chibchan and (Proto-)Tupí–Guaraní kinship terms is the existence of forms with and without an initial alveodental stop: compare the two Proto-Chibchan forms reconstructed here for ‘brother-in-law’, *u^{m}ba and *^{n}du^{m}ba, with Proto-Tupían *up ‘father’, which has a reflex *-up* in Awetí and a reflex *t-uβ in Proto-Tupí–Guaraní (Rodrigues 2010: 8). In some cases presented in table 112, in elements encoding grammatical information and also in the terms for ‘woman’s breast’ or ‘louse’, similar forms are also found in Macro-Jê languages (see below, subsection 3.3), and it would be difficult to claim that these elements are indicative of an exclusive relationship between Chibchan and Tupían.

[92] Andrey Nikulin (p.c.) reconstructs the following Proto-Tupían forms: *ka dative/allative; *ʔək ‘manioc’; *kũ ‘tongue’; *ŋãm ‘breast’; *ŋɨp ‘louse’; *pap ‘to die’; *pe punctual locative, dative; *atuk ‘to bathe’; *tup ‘to see’.

[93] For a similar sound correspondence across language boundaries, compare, for instance, Proto-Macro-Jê *kɤ, Proto-Tupían *pe and Proto-Cariban *pi ‘skin’ (cf. Nikulin 2015a: 94).

As mentioned above, Herzog (1886) also saw some parallels between Central American Chibchan and Cariban languages.[94] Comparing Chibchan and Cariban, or rather the two proto-languages, a few parallels can indeed be found, as shown in table 113. Again, these coinciding forms are far from sufficient in order to postulate a close genealogical connection between both language families – however, very little Proto-Cariban vocabulary has been published so far; the only source at my disposal was Gildea and Payne (2007).

TABLE 113
SOME PROTO-CHIBCHAN–PROTO-CARIBAN LOOKALIKES

Set	Proto-Cariban	Proto-Chibchan
1	*a-t**ɨpô** ‘seed’	*u^{m}ba ‘face, eye, fruit’
2	*tiri**kô** ‘star’, ?*piti**kô** ‘small’	*k^{w}a ‘seed, fruit’
3	*(t)**wô**(nô) ‘to kill, shoot’	*^{ŋ}gua ‘to kill’
4	?***pôre**(-pɨ/pa) ‘leg’	*ka^{n}d- ~ *kat- ‘stick, bone, tree’
5	*a**pô**-rɨ ‘arm’	*ka ‘leaf’
6	*j**ar**e ‘leaf’	*ha^{n}d- ~ *hat- ‘hand’
8	*a**kôrô** ‘dog’	*ku(^{n}d-) ‘feline$_2$’
9	*a**rokɨ** ‘penis, tail’	*^{n}duh(k) ‘tail’
10	***(j)ô**-rɨ ‘tooth’	*^{n}duʔ ‘tooth’

Proto-Cariban reconstructions from Gildea and Payne (2007); Proto-Chibchan reconstructions are mine.

Other parallels do exist – concerning the domain of bound morphology, compare the following example (18) from Katxuyana, a Cariban language from the Brazil–Suriname border region.

Katxuyana (Gildea 2015: 3)
(18) *n-**ot**-woni-wɨ*
3-**REFL**-kill-PST
‘He killed himself.’

[94] Likewise, Weißhar (1982: 356) hypothesized a certain connection between Chibchan and Cariban in the conclusions of his work on the external relations of Yanomaman and Warao. (“Im Gegensatz zu VOEGELIN und GREENBERG, die das Caribische und Chibcha zwei unterschiedlichen Phylen zuordneten, gehen wir vielmehr von einem Block aus, der u.a. das Caribische und Chibcha umfaßt, dem aber auch das Yanomama und Warao angehören.”)

Katxuyana *ot-* has cognate counterparts in other Cariban languages (cf. Gildea 1998; Meira et al. 2010: 505–12), derives from Pre-Proto-Cariban reciprocal *ôte- (cf. Meira et al. 2010: 513), and strikingly resembles Proto-Chibchan valency-reducing *a^{n}d- in form and function. Also, both the Proto-Cariban and the Proto-Chibchan valency-reducing elements in question are prefixes and thus coincide in position.[95]

There are some additional coincidences between Chibchan, Cariban and Tupían languages. One example is the Proto-Chibchan term for the agouti, *ˈkuri 'paca (*Agouti paca*)', reconstructed as such by Constenla Umaña (1981: 405) on the basis of some Central American Chibchan languages, which is most probably related to Proto-Cariban *akuri 'agouti' (cf. Meira & Franchetto 2005: 139)[96] and Proto-Tupían *akutʔi 'agouti' (cf. Silva 2007: 229) (and, of course, to Proto-Emberá *khuɾiˈhia 'paca, agouti', see subsection 3.2.3.2).

As such, the parallels uncovered so far between Proto-Chibchan, Proto-Cariban, and Proto-Tupían are probably not enough to substantiate a close genealogical relationship. On the other hand, they are definitely not so scanty that we can readily dismiss Herzog's hypothesis of a genealogical link between Chibchan, Tupían (note that Herzog himself only mentions Tupí–Guaraní languages, not Tupían languages), and Cariban.

The fact that Herzog (1886) saw a genealogical connection between Chibchan languages of Central America with both Cariban and Tupían implies that he was aware of a genealogical link between Cariban and Tupí–Guaraní languages. This is remarkable insofar as a Tupían-Cariban connection was explicitly put forward much later by Rodrigues within his Tu-Ka-Jê proposal (2000, 2009).[97] Rodrigues' Tu-Ka-Jê additionally includes Macro-Jê languages. A possible connection between Chibchan and Macro-Jê languages will be discussed in the following subsection.

[95] Andrey Nikulin (p.c.) proposes that the Katxuyana reflexive morpheme in question goes back to Proto-Cariban *tə- reflexive, with potential cognates in Proto-Tupían (*tə- reflexive), in Proto-Macro-Jê (*ta-, labelled 3.COREF) and Boróro. Indeed, these latter reflexive morphemes are reminiscent of the Kuna reflexive prefix *na-* which encodes the meaning 'own' and notions of reflexivity according to Holmer (1947: 95, 125–6), of Kuna *te-* 'this (the same)' (Holmer 1946: 190), and of the Guatuso prefix *ɾa-* ~ *ɾi-* which encodes notions of reciprocity or reflexivity (Constenla Umaña 1998: 126).

[96] I am grateful to Matthias Urban (p.c.) for pointing out this possibility.

[97] A more detailed discussion of some morphological evidence for a Cariban–Tupían connection is provided by Meira et al. (2010: 512–5).

3.3 An alternative proposal: Chibchan and Macro-Jê

In this section I will argue that there is evidence for a genealogical connection between Chibchan and the Macro-Jê languages. The latter are spoken in eastern South America and comprise the Jê and Jabutí languages, Karajá, and Chiquitano, among others (more details on Macro-Jê will be given below). My argument is essentially based on parallels between Chibchan grammatical forms and their counterparts in the abovementioned Macro-Jê languages. The parallels in question are summarized in table 114. They include

- three sets of shared allomorphs:
    - two suppletive forms of the verb 'to go' (sets 2 and 3)
    - three valency-reducing prefixes (sets 4 to 6)
    - two relational elements (sets 14 and 15)
- formally and functionally similar grammatical elements derived from a root 'to do, make' (set 1)
- a shared preverbal element encoding a generic object (set 7)
- shared alternations of thematic vowels (set 8)
- a shared derivational element relating to shape (set 13)
- shared sets of
    - four case-markers (sets 9 to 12)
    - two deictic elements (sets 16 and 17).

TABLE 114
(PROTO-)CHIBCHAN–(PROTO-)(MACRO-)JÊ COGNATE SETS IN MORPHOLOGY

Set	Function	(Proto-)Chibchan	Proto-Macro-Jê	Proto-Jê, Proto-Northern Jê, Proto-Cerrado Jê (= Proto-Amazonian Jê)	Single Macro-Jê languages
1	PC, PNJ ‘to do, make’ (further grammaticalization into a causative prefix, instrumental postposition)	*õ		PNJ *ɔ	
2	PC ‘to go’, PMJ, PCJ ‘to go, come (pl.)’ (further grammaticalization into a marker of future tense, purposive / dative / allative)	*^{m}ba	*mɔ̃ŋ (Na 55)	PCJ *mɔ̃ (Nb 297)	
3	PC ‘to go’, PMJ, PCJ ‘to go, come (sg.)’ (further grammaticalization into a marker of future tense, dative/allative)	*^{n}daʔ ~ *taʔ	*tɛ̃(C) (Na 55); *jɛ ~ *tɛ	PCJ *tɛ̃ (Nb 299)	*ʧɛ*- Arikapú (Jabutí language) direct-object marker (RV 529)

4	valency-reducer$_1$	*a-	*a-		Apinajé *a-* (Ol 131), Karajá *a-* (R 88)
5	valency-reducer$_2$	*aʔ-			Apinajé *aʔ-* (Ol 131)
6	valency-reducer	*a^{n}d-			Apinajé *at-* (Ol 131)
7	(body), generic object prefix	*apa			Apinajé *ap ~ aw* (Ol 131),[98] *ɔ-* Karajá antipassive (R 54)
8		Default form of the verb stem ends in *a, verb stems of certain other forms end in a non-predictable, more fronted/higher vowels in Bribri and Cabécar			Default (finite) form of the verb stem ends in *a*, verb stems of certain other forms end in a non-predictable, more fronted/higher vowels (Jê, Nikulin 2016b) – for instance in Apinajé (Ol 194) or in Pykobjê (Am 109–14)
9	locative	*ki	*ki		Xoklény *ki* (H 202); Karajá *kɪ* (R 61)
10	PC allative/dative; PJ indirect-object prefix	*ka		*ka	Apinajé *ka-*, Kaingáng *kã-* (RV 552)
11	PC locative, PMJ locative/instrumental	*^{n}da ~ *ta	*jɛ ~ *tɛ	*jɛ ~ *tɛ	*tɛ* Apinajé ergative (Ol 178); *ʧɛ Proto-Jabutí locative

[98] This form is reminiscent of Proto-Tupí–Guaraní *aβa ‘man, human being’ (Mello 2000: 150).

12	instrumental/locative	*^{n}di	*ji ~ *ɾi		=*dɪ* Karajá instrumental (R 102); -*ɾi* Apãniêkrá locative (Al 87)[99]
13	derivational element ‘three dimensions’	*-k^{w}a	*-kua	*-kua (cf. RV 558)	*-ku, *-ko Proto-Jabutí (cf. RV 557–8, 569)
14	relational element	*^{n}d(i)-	*j-, *ɲ- (Na 7)	*j- (RV 551)	
15	PC relational element; PJ 3rd person	*ʦ(i)-	*ʃ- (Na 7)	*s- (RV 551)	
16	PC proximate demonstrative; PMJ, PJ 1st person	*hi ~ *iʔ	*ij		Pre-Chiquitano 1st person singular markers *iʂ- feminine, and *ij- masculine (Nikulin forthc.)
17	PC distal demonstrative; PMJ, PJ 2nd person	*a	*a	*a- (RV 551)	Pre-Chiquitano *a- 2nd person singular marker (Ad 13)

My own reconstructions, if not indicated otherwise.

[99] Willem Adelaar (p.c.) observes a similarity with the Guaraní comitative marker /-ni/ [ndi].

These parallels will be discussed in more detail in subsection 3.3.2 to 3.3.4. They constitute a central argument in support of a genealogical connection between Chibchan and Macro-Jê languages. This hypothesis is further corroborated by systematic sound correspondences which are also reflected in several basic lexical items (including shared homophones, cf. subsection 3.3.5). It is likewise in line with the similar typological profile of Proto-Chibchan and Proto(-Macro-)Jê that will be discussed in subsection 3.3.6.

So far, a genealogical connection between Chibchan and Macro-Jê languages has not been proposed. Nonetheless, Herzog's (1886) and Rodrigues' (2000, 2009) proposals, briefly mentioned in subsection 3.2.3.4 are, to a certain extent, indirect predecessors, given that these investigators propose a genealogical connection between Chibchan, Tupí–Guaraní, and Cariban (Herzog) and between Tupían, Cariban, and Macro-Jê (Rodrigues). Also, Rivet (1953) identified parallels between Mashubí on the one hand (Arikapú, one of the two Jabutí languages, cf. Caspar 1955, cited in Ribeiro & van der Voort 2010: 520), and, on the other hand, Chibchan and other languages that he classified as such, among which are also Páez and Andakí. This, however, does not imply that Rivet would have argued for any connection between Chibchan proper, as defined by Constenla Umaña (1981), and Macro-Jê, as defined by Nikulin (2015a) or Ribeiro and van der Voort (2010). [100] Finally, Loukotka (1948: 197) suggests that it would be interesting to compare Jê and Chibchan languages, without giving further details.[101]

Chibchan and Macro-Jê are compared in this thesis by a juxtaposition of proto-language forms and/or a juxtaposition of modern forms. Given the fact that relatively little reconstructed Macro-Jê materials are available so far (an exception is Nikulin 2015a), and given the conservative character of most Northern Jê languages (see below), I will mostly base my argumentation on parallels between Chibchan with this subgroup of Macro-Jê.

Admittedly, chance similarities and borrowing cannot be excluded in every single instance, but they cannot account for all the parallels in Chibchan and Macro-Jê

[100] In fact, Rivet (1924a) was very sceptical *vis-à-vis* the existence of a linguistic unit he calls "Žę", comprising Southern Jê (e.g., Kaingáng, Xokléng), Central Jê (Xerénte), Northern Jê (e.g. Timbira, Kayapó), Jeikó, Borum, Kamakã, Maxakalí, Pataxó, Purí, and Ofayé ("Šaváante Opaié"). He states that this is "the most articifially constituted of all South American families, the *caput mortuum* of South American linguistics" (Rivet 1924a: 697–700). ("Cette famille est, de toutes les familles sud-américaines, celle qui a été le plus artificiellement constituée. Elle est comme le *caput mortuum* de la linguistique sud-américaine. Sa révision soigneuse et complète s'impose, sur des bases vraiment scientifiques." ibid.: 697)

[101] "Dans le Brésil central, on n'a pas encore éclairci les rapports du groupe Žé avec les tribus voisines et, ce qui est plus intéressant, avec les langues du groupe Čibča." (Loukotka 1948: 197). Note, however, that Loukotka (ibid.: 196) was among the proponents of an original homeland of Chibchan-speaking groups in Central America, not in South America.

morphology discussed here, which include several sets of morphemes and several cases of shared allomorphy. Additionally, recent borrowing between Chibchan and Macro-Jê languages can largely be ruled out as an explanation, given the geographically distant distribution areas – at present, Macro-Jê languages are all spoken south of the Amazon.

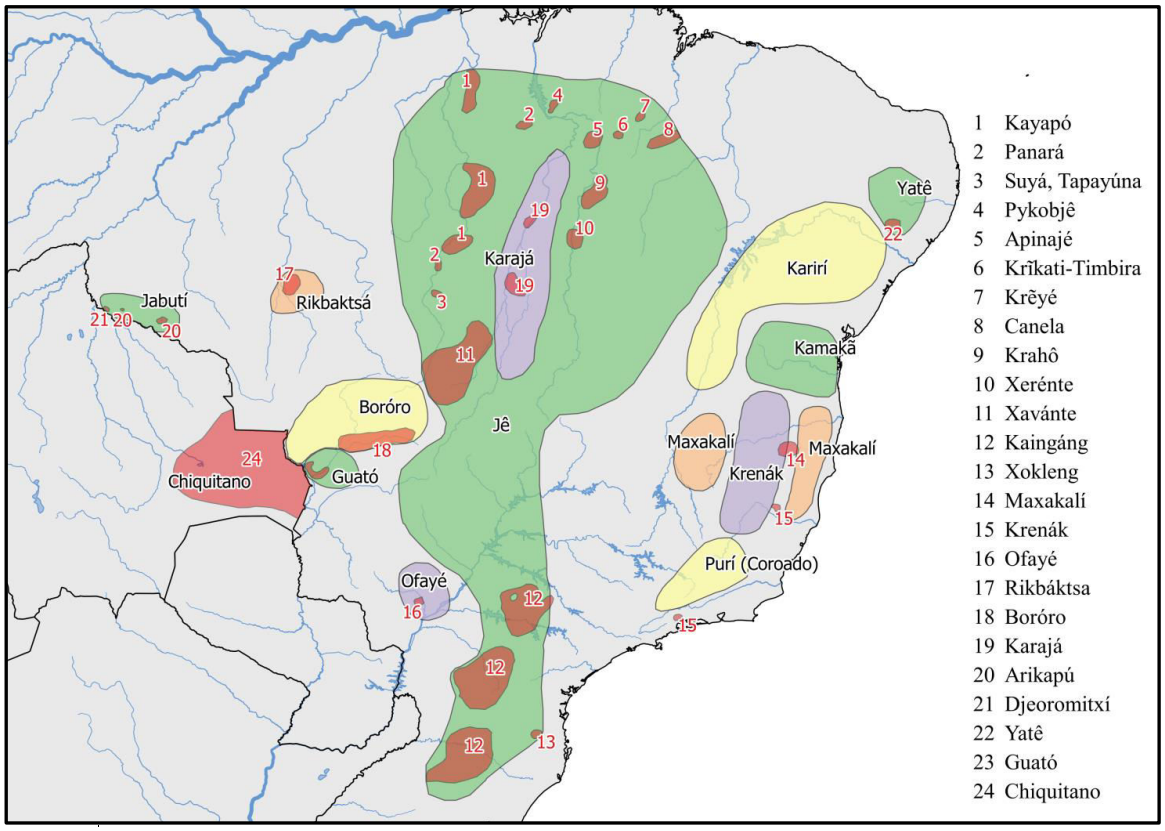

Figure 4: The distribution of languages that have been classified as Macro-Jê. The map was created by Arjan Mossel, based on data from Ribeiro & van der Voort (2010: 546) and Simons and Fennig (2017). Chiquitano is mainly spoken in Bolivia, but also in some adjacent areas of Brazil.

The structure of this subsection is as follows

- Subsection 3.3.1 briefly introduces the Macro-Jê language family.
- Subsection 3.3.2 presents and discusses the verb 'to do, make' and shared allomorphs of the verb 'to go' and parallel grammaticalization chains in Chibchan and Macro-Jê languages, which obtain for these forms (compare 1 to 3 in table 114).

- Subsection 3.3.3 presents and discusses parallels in verbal morphology, including shared allomorphic variation in the domain of verbal valency (compare 4 to 8 in table 114).
- Subsection 3.3.4 presents and discusses Chibchan–Macro-Jê parallels in nominal morphology, among which are shared sets of case-marking postpositions, possessive markers and deictic elements (compare 9 to 17 in table 114).
- Subsection 3.3.5 discusses some recurrent Chibchan–Macro-Jê sound correspondences.
- Subsection 3.3.6 discusses some typological parallels of (Proto-)Chibchan and (Proto-)Macro-Jê.

3.3.1 Macro-Jê languages

The term 'Macro-Jê' was first introduced by Mason (1950: 287). At present, several authors (Rodrigues 1999; Adelaar 2008; Ribeiro & van der Voort 2010; Ribeiro 2012: 265, Nikulin 2015a: 10) acknowledge the existence of a Macro-Jê stock, one of the few big genealogical units of eastern South America. The Macro-Jê status of several languages is still hypothetical. The following languages have been included so far in this stock by different authors (see figure 4): Southern, Central, and Northern Jê, Borum (Krenák or Botocudo), Boróro, Chiquitano (not accepted as Macro-Jê by Rodrigues 1999), Jabutí languages, Kamakã, Karajá, Karirí, Maxakalí, Ofayé, Purí, and Rikbaktsá (cf. Ribeiro & van der Voort 2010: 548; Adelaar 2008). Different opinions do exist, for instance, concerning the Macro-Jê status of Guató and Yatê (Fulniô). They were classified as Macro-Jê by Greenberg (1987) and Rodrigues (1999), whereas other authors are more careful concerning a Macro-Jê status of these languages (e.g., Ribeiro and van der Voort 2010: 546–8). In line with that, Macro-Jê as a whole is often treated as a hypothetical unit or stock (e.g., by Rodrigues 1999; Ribeiro & van der Voort 2010), but this hypothetical character may sometimes also be based on a general lack of research. Little reconstructive work on Proto-Macro-Jê has been undertaken so far (an exception is Nikulin 2015a: 10–84). Within Greenberg's 1987 work, *Language in the Americas*, the Macro-Jê subgroup of alleged Amerind is exceptional in that it has received support in several cases by subsequent research (Adelaar 2013: 116). A further step ahead, Rodrigues' (2000, 2009) Tu-Ka-Jê unit, based on shared allomorphy, is the only macro-level proposal in South America that has met with a favorable reception so far (e.g., Nikulin 2015a), possibly alongside Viegas Barros' (2005b) Macro-Jê–Macro-Guaicurú proposal (cf. Nikulin & Carvalho 2018).

In the present work, I follow Jolkesky's (2016: 259) careful approach, only including data from Jê proper, Karajá, Jabutí languages, and Chiquitano as nuclear

Macro-Jê of sorts. Having said that, I do not want to suggest that none of the other languages belongs to the Macro-Jê stock (see above). Instead, my decision is based on the fact that the genealogical connection between Jê, Karajá, Jabutí, and Chiquitano has been duly demonstrated in recent work of specialists (Ribeiro & van der Voort 2010; Adelaar 2008; Ribeiro 2012; Nikulin 2015a). In what follows, I will briefly present these four language groups.

3.3.1.1 Jê

Jê languages are spoken in central-eastern Brazil. The term "Gentis Gès" was probably first used by Martius (1867: 134) to refer to the Kaiapó, Apinajé, Krahô, Xavánte, and Xerénte, among others. This implies that Martius also recognized the genealogical connection between Central Jê (Xavánte, Xerénte) and Northern Jê languages (cf. ibid.: 134–66). The shared origin of Kaingáng (Southern Jê) and the other Jê languages became clear only later. Among the first researchers who were aware of this connection was probably Nimuendajú (1914: 376). At present, Rodrigues (2012: 269) proposes a threefold distinction of Northern, Central, and Southern Jê. Ribeiro and van der Voort (2010: 547) and Nikulin (2015b: 276; 2016a) instead propose a major division between Southern Jê on the one hand and Proto-Amazonian or Proto-Cerrado Jê on the other. This subgroup is then further divided into two branches, namely Northern Jê and Central Jê (see figure 5), by Ribeiro and van der Voort (2010) and Nikulin (2016a).

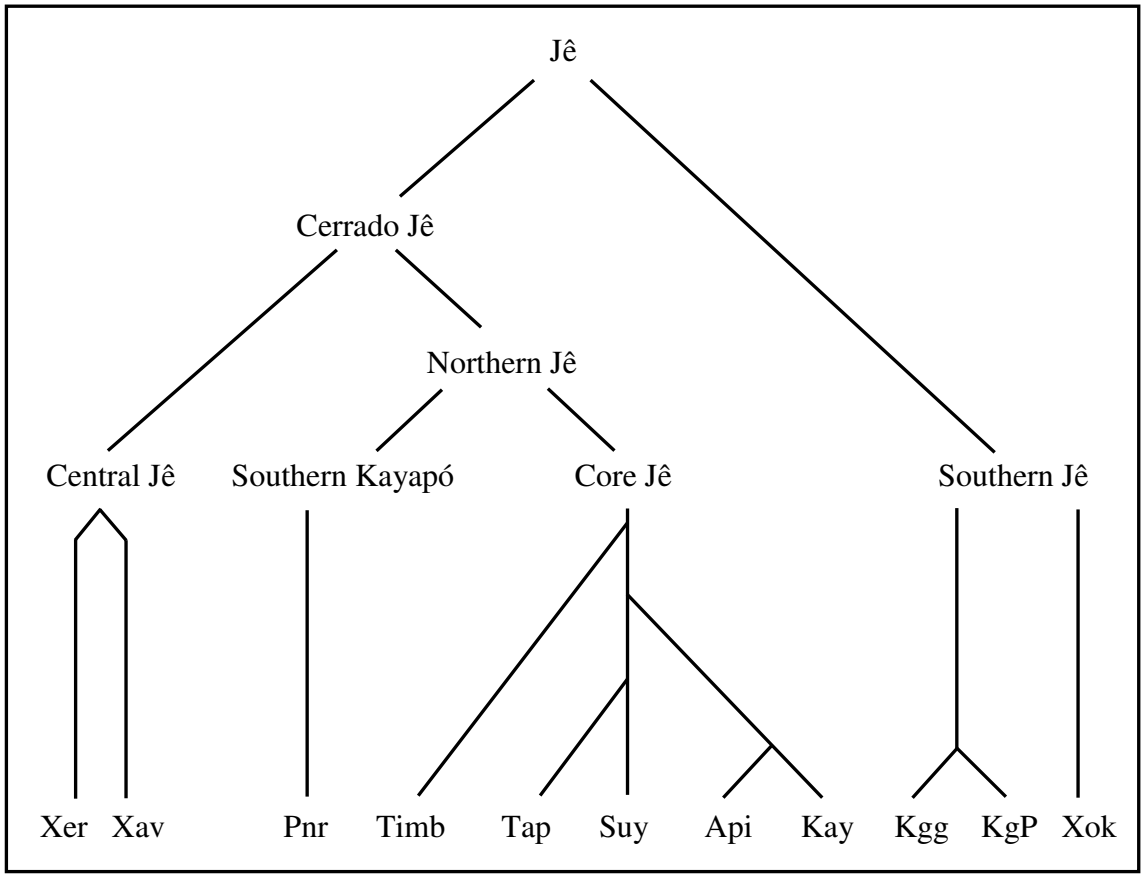

Figure 5: Phylogenetic structure of the Jê family (Nikulin 2016a, 2017). Abbreviations: Xer = Xerénte; Xav = Xavánte; Pnr = Panará; Timb = Timbira (including Apãniêkrá and Krahó); Tap = Tapayúna; Suy = Suyá (Kĩsêdjê); Api = Apinajé; Kay = Kayapó (or Mẽbêngôkre, including Xikrín); Kgg = Kaingáng; KgP = São Paulo Kaingáng; Xok = Xokléng.

Within Jê proper, Northern Jê languages are conservative insofar as they did not undergo extensive vowel shifts like Central and Southern Jê languages (Ribeiro & van der Voort 2010: 554). This makes this subgroup particularly relevant and suitable for comparative research. Another argument in favor of a relatively conservative character of Northern Jê lexicon is that in this domain, Chiquitano shows more similarities with Northern Jê than with Southern Jê (cf. Adelaar 2008: 15). Within Northern Jê, Apinajé, Kayapó (or Mẽbêngôkre, including Xikrín), Suyá (Kĩsêdjê) and Timbira (including Apãniêkrá, Krahô, etc.) share several morphosyntactic peculiarities, for instance in ergative marking and some non-finite verb forms (Castro Alves 2004: 148). Panará stands out among the Northern Jê languages for its relatively complex morphology and some typological peculiarities, for example, in the domain of word order (Oliveira 2005: 27). The morphological complexities of Panará are innovations (Ribeiro and van der Voort 2010: 552).

Aikhenvald (2015: 48) gives a brief overview of some non-linguistic features shared by peoples speaking Jê languages: their traditional habitat is the savannah, not the rainforest. As to cultural elements and features, traditionally, there is no use of hammocks nor of canoes among Jê-speaking groups, no pottery, and slash-and-burn agriculture is not traditionally practiced either. The structure of villages is circular, consisting of a plaza which is surrounded by individual, circular-roofed houses.

3.3.1.2 Karajá

Karajá is a Macro-Jê language with four dialects, Karajá proper, subdivided into Southern Karajá and Northern Karajá, Javaé, and Xambioá. Karajá is spoken by some 3,000 people in the Goiás, Tocantins, Pará, and Mato Grosso states of Brazil (Ribeiro 2012: 1). It was first argued to be genealogically related to Northern and Central Jê languages by Steinen (1886: 316). This proposal was later confirmed by Davis (1968) and Ribeiro (2012: 262–87). The Karajá area of distribution is surrounded by speakers of Jê languages. Verbal morphology in Karajá is more complex than in most Jê and Jabutí languages, in the sense that many categories are indicated by bound morphology in this language (cf. Ribeiro 2012: 38). These morphological complexities are not retentions from the proto-language but innovations (Ribeiro and van der Voort 2010: 552–3). A central feature of Karajá is the correspondence of *k* in female speech with *Ø* in male speech (Ribeiro 2012: 165–6).

3.3.1.3 Jabutí languages

Jabutí languages consist of Djeoromitxí (Jabutí) and Arikapú. Djeoromitxí still had some fifty speakers at the beginning of the twenty-first century, whereas there was only one speaker of Arikapú at that time. Jabutí peoples inhabit the area of the Rio Branco, which is a right tributary of the Guaporé river. The latter separates Beni, a northeastern province of Bolivia, from Rondônia, a state of western Brazil (Ribeiro & van der Voort 2010: 518). The genealogical connection of Jabutí languages with Jê languages was first explicitly mentioned by Nimuendajú in a private correspondence from 1935 (Ribeiro & van der Voort 2010: 521). Apparently independently, the Jabutí family was also identified as belonging to the Macro-Jê stock by Greenberg (1987: 66).[102] This genealogical connection was finally demonstrated by Ribeiro and van der Voort (2010).

[102] According to Ribeiro and van der Voort (2010: 521) "[t]he very fact that Greenberg (1987) adopted this classification [of Jabutí languages as Macro-Jê] may have been a reason that it did not become broadly accepted among Americanists."

3.3.1.4 Chiquitano

Chiquitano is an endangered language, spoken in the Bolivian department of Santa Cruz and adjacent areas in Brazil (state of Mato Grosso). There are several different dialects, such as Bésɨro (Lomeriano), Ignaciano and Migueleño (Nikulin, forthc.). Estimations of speaker numbers vary (between 1,000 and 6,000 for Bésɨro) (Sans 2010: 31–2). A description of the Bésɨro dialect has recently been prepared by Sans (2013). A Macro-Jê status of Chiquitano was first proposed by Greenberg (1987: 66) and later confirmed by Kaufman (1990) and Santana (2006). The validity of a genealogical connection with Jê languages and Karajá was demonstrated by Adelaar (2008), and the Macro-Jê status of this language has been accepted by several experts of this language family (e.g., Ribeiro 2012: 263; Jolkesky 2016: 259–64). The Macro-Jê status of Chiquitano may be illustrated in table 115 by the juxtaposition of some Proto-Jê reconstructions with Chiquitano data published by Gilij (1782: 357–63). Gilij's forms were chosen here because they are relatively free of affixed elements, except for a final element <-s ~ -z>, which indicates absence of a speech-act participant possessor (Willem Adelaar, p.c.).

TABLE 115
CHIQUITANO–PROTO-JÊ LEXICAL PARALLELS

Set	English	Chiquitano	Proto-Jê
1	arm	<i**pá**s>[103]	*pa (RV 558; Nb 287)
2	tongue	<**otù**s>	*j-õtɔ (RV 559), *çõj̃tʌ/*ɲõj̃tʌ (Nb 292)
3	rain	<**taà**s>	Proto-Central Jê *tə̃ (Na 24)
4	stone	<**caà**s>	*kɛn (RV 560)
5	fire/f-wood	<**peè**z>	*pĩ (RV 550), *pĩm (Nb 299)
6	hand	<**eê**s>	*j-i- (RV 552), *çĩp=kɾa / ɲĩp=kɾa (Nb 287)
7	meat	<a**ñe**z>	*j-ĩ (RV 558); *ɲĩ (Nb 299)
8	nose	<**iña**s>	*j-ĩja (RV 569), *ʃ̃iya(C) / *ɲĩya(C) (Na 71)
9	leg/bone	<**ii**s>	*s**i** (RV 558); *ç**i**/*j**i** 'bone' (Nb 296)
10	tooth	<**oò**s>	*j-ua 'tooth' (RV 565); *ç**ɔ**/*j**ɔ** (Nb 289)
11	lip/mouth	<**arù**s>	*j-ar- (RV 552); *ç**at**ᵊ-kɔ / *j**at**ᵊ-kɔ (Nb 287)

Sources are Gilij (1782: 357–63) for Chiquitano; Ribeiro and van der Voort (2010); Nikulin (2015a, 2015b) for Proto-Jê.

[103] The prefix *i-* indicates a 3rd-person possessor (cf., e.g., Nikulin forthc.).

Like in Karajá, there is also a distinction between male and female speech in Chiquitano (e.g., Nikulin forthc.). Also, Chiquitano is morphologically more complex than most Jê and Jabutí languages. The complexity of Chiquitano largely resides in the irregularities of its morphology (Adelaar 2008: 11) and is probably an innovation rather than a retention (Willem Adelaar, p.c.).

3.3.2 Shared allomorphs and parallel grammaticalization chains

This subsection presents and discusses the formal coincidence of Chibchan and Northern Jê 'to do, make', and of the Chibchan and Macro-Jê suppletive forms of the verb 'to go', and the parallel grammaticalization paths of these verbs in Chibchan and Macro-Jê, halfway between verbal and nominal morphology. It remains to be investigated to what extent the parallel grammaticalization paths described below are triggered by the similar typological profiles of both language groups (see below, subsection 3.3.6), by universal tendencies of grammaticalization, or by a specific "drift" in the shared ancestor language which continued in the daughter languages (cf. Sapir 1921: 184). In any case, the three forms dealt with below imply only one cognation set each, given that the verbal roots and the grammatical elements derived from them are etymologically related.

3.3.2.1 *õ and *ɔ 'to do, make'

Both in Chibchan and Northern Jê languages, a verb 'to do, make', a causative proclitic, and an instrumental/comitative postposition all seem to derive from a proto-form *õ (Proto-Chibchan) or Proto-Northern Jê *ɔ (< Proto-Macro-Jê *ʌ, according to Nikulin 2015a: 41) (compare set 1 in table 114 above). Table 116 illustrates the reflexes of this element and their uses in Kuna, Rama (both Chibchan) and in Mẽbêngôkre and Apãniêkrá (both Northern Jê).

TABLE 116
VERBS 'TO DO, MAKE' AND DERIVED ELEMENTS IN TWO CHIBCHAN AND NORTHERN JÊ LANGUAGES

	Kuna	Rama	Mẽbêngôkre	Apãniêkrá
'to do, make'	—	*uŋ*	*ɔ*	*tɔ*
Preverbal causative marker	*o-*	*ju=* comitative, instrumental	*ɔ=*	*tɔ=*
Instrumental/comitative marker	(*-jo*)	*u*	*=ɔ*	*=tɔ*

Sources are Holmer (1947) (Kuna), Craig (1989) (Rama), Reis Silva (2003) (Mẽbêngôkre) and Castro Alves (2004) (Apãniêkrá).

The reflexes of *õ and *ɔ are sometimes combined with an element that precedes it: in Kuna and Rama, it is *j-*, in Apãniêkrá, it is *t-*. Rama *j-* and Apãniêkrá *t-* are not part of the morpheme originally (cf. Craig 1989: 126; Oliveira 2005: 415). However, Chibchan *j-* and Apãniêkrá *t-* are not necessarily cognate.

Chibchan

In what follows, I briefly present evidence from Chibchan, where a verb *õ 'to do, make' became grammaticalized as a preverbal causative marker and as an instrumental/comitative postposition.

The verb 'to do, make'

I reconstruct a Proto-Chibchan verb *õ 'to do, make' on the basis of Boruca *óŋ* 'to do, make' (Quesada Pacheco & Rojas Chaves 1999: 150), Cabécar *ő́-ɾ* (the remotospective perfective form of *wã́* 'to do, make') (Margery Peña 1989: 449), and Rama *u:ŋ* 'to make' (Craig 1989: 207). The use of the Rama form is illustrated in example (19), in the context of the expression 'to be X years old', literally 'have made X years':

Rama (Craig 1989: 90)

(19) *ti:skam* *naiŋ* *jiɾ* *paŋsak* *j-**u:ŋ**-u*
child this year three 3-**make**-PFV
'This child is three years old.'

Preverbal causative marker

A verb 'to do' has been observed as a source of causative markers across different languages (e.g., Heine & Kuteva 2004: 117–8). Indeed, in Cabécar, for instance, the verb 'to do, make' conveys a causative meaning if it follows the verb (cf. Margery Peña 1989: 449–50). Instead, a preverbal causative marker derived from *o or *õ is reflected in Kuna *o-takka* 'to show, awaken', from *takka* 'to see' (Holmer 1947: 121). Kuna *o-* is cognate with the causative marker *u-* in Kogi ***u****nafi* 'to bring' versus *lafi* 'to come' (cf. Olaya Perdomo 2000: 781), according to Holmer (1947: 28), and probably also with Pech *ũ̀-* in ***ũ̀****ː-wìʔ-t-* 'to dance with', literally 'to cause to dance' (cf. Holt 1999a: 61; Adelaar and Muysken 2004: 73). In Rama, a preverbal element *ju=* likewise has a comitative or instrumental meaning (Craig 1989: 115).

Instrumental postposition

It is easily conceivable that in a language with SOV as a basic word order, a preverbal causative marker *o or *õ was reassigned to postnominal position, encoding comitative or instrumental notions. Indeed, this is exactly what seems to have happened in several Chibchan languages. In Rama, there is a comitative/instrumental postposition *u* (Craig 1989: 115). The instrumental use of the cognate instrumental/comitative postposition *ju* in Guatuso is illustrated in (20):

Guatuso (Constenla Umaña 1998: 112)

(20) *puˈɾuto* *ˈju* *Ø-ri-xaˈti-je*
axe **with** 3.OBJ-3.AG-cut-REA
'S/he cut it with the axe.'

Kuna *-jo* indicates "the state of being one of a certain group" (Holmer 1947: 58) and may belong here, too.

Northern Jê

In northern Jê languages it is possible to identify forms and grammaticalization paths which are most similar to those observed in Chibchan above.

The verb 'to do, make'

In Mẽbêngôkre, a northern Jê language, the use of a verb *ɔ* 'to do' is illustrated in the following example (21):

Mẽbêngôkre (Reis Silva 2003: 44)

(21) *mɤj nẽ ga ɔ ɲũ̃*
what FOC 2 **do** sit
'What you are doing sitting down?'

Preverbal causative marker

In preverbal position, the Mẽbêngôkre verb 'to do, make', shown in example (21), may also encode a causative meaning (cf. Reis Silva 2003: 46). Similar functions are also found in other Northern Jê languages: The following example from Apinajé, a Northern Jê language, illustrates the use of the corresponding proclitic *ɔ*= (22):

Apinajé (Oliveira 2005: 265)

(22) *na ka kawə ɔ=dɔt*
REA 2 basket **make**=full
'You filled the basket.'

According to Oliveira (2005: 265), no intervening element is allowed between *ɔ*= 'make' and *dɔt* 'full' in (22); nonetheless, *ɔ* cannot be interpreted as a prefix because of stress patterns.

In some Jê languages, an element *t-* is combined with the reflex of *ɔ 'to do, make', for instance, in Apãniêkrá *tɔ* 'to do, make' (cf. Castro Alves 2004: 73), or in the causative proclitic of the same language, *tɔ*= (cf. ibid.: 76). A tentative explanation of this element *t-* will be given below.

Instrumental postposition

How a preverbal causative marker could be reassigned the status of an instrumental/comitative postposition is illustrated by the following example from Apinajé (23):

Apinajé (Oliveira 2005: 160)

(23) *pa kɔt paj Zɛbaketi ɔ grɛ ɾatʃ kumrɛtʃ*
1 FUT 1.FUT Zé Cabere **COM** dance much INT
'It's me that I want to dance all night with Zé Cabere.'

In this example, *Zɛbaketi ɔ grɛ* can be analyzed either as 'Zé Cabere with dance' or as 'Zé Cabere make dance'. A similar, ambivalent case is attested in Mẽbêngôkre 'I bring meat to the village' (24):

Mẽbêngôkre (Reis Silva 2003: 46)

(24) *krĩ* *mɤ̃* *nẽ* *ba* *mrɯɲi* **ɔ** *tẽ*
village ALL FOC 1.AG meat **make** go
'I bring meat to the village.'

In this example, one may analyze *mrɯɲi ɔ tẽ* either as 'meat with go' or 'meat make go'. Thus, the verb 'to do, make' seems to have been reinterpreted as a causative marker, and, possibly in a further step, it has been reassigned to post-nominal position in Northern Jê languages, exactly as in certain Chibchan languages.

3.3.2.2 *^{m}ba 'to go$_1$'

In Chibchan and probably also in Macro-Jê languages, a verb 'to go', a future tense marker, and a dative / allative / purposive marker derive from *^{m}ba, *^{m}bã or a related form (compare set 2 in table 114 above). These developments are in line with the following grammaticalization paths that have been observed in other language groups across the world:

(1) 'go' > future tense marker (Heine & Kuteva 2004: 161–3)
(2) 'go' > purpose (Heine & Kuteva 2004: 163–5)
(3) 'go' > allative (Heine & Kuteva 2004: 160–1)
(4) allative > purpose (Heine & Kuteva 2004: 398–40)
(5) allative > dative (Heine & Kuteva 2004: 37–8)

As to Chibchan and Macro-Jê languages, reflexes of the forms in question are presented in table 117 for Rama (Chibchan), Djeoromitxí (Jabutí language, Macro-Jê) and Apinajé (Jê proper).

TABLE 117
POSTPOSITIONS AND PROCLITICS/PREFIXES DERIVED FROM A VERB 'TO GO' IN RAMA AND TWO MACRO-JÊ LANGUAGES

	Rama	Djeoromitxí (Jabutí language family)	Apinajé
'to go'	*maŋ* ~ *baŋ* 'go!' (170, 255)		*mõ* 'to go (wandering)' (129)
future tense /prospective aspect	*-baŋ* (170)	*ma* (78)	
goal, purposive/allative, dative	*baŋ* goal, purpose (114)	*ma* dative, 'for' (116)	*mə̃* allative (54)
Preverbal goal, purposive/dative case-marker	*ba=* goal, purpose (114)		*mə̃-* dative (134)

Sources are Craig (1989) (Rama), Pires (1992) (Djeoromitxí), and Oliveira (2005) (Apinajé). Page numbers in parentheses.

In some Chibchan and Jê languages, the dative-marking element has been assigned a preverbal position. The corresponding grammaticalization chains in Rama have been described in detail by Craig (1989: 170, 252).

Chibchan

In what follows, I briefly present evidence for a verb 'to go' that became grammaticalized as a future tense marker, a purposive/benefactive postposition, and a preverbal dative marker in several Chibchan languages.

A verb 'go'

It is possible to reconstruct an element *^{m}b- 'to go' for Proto-Chibchan, followed by a vowel *a or *i. Proto-Chibchan *^{m}ba (or *^{m}bã) 'to go' has reflexes, for instance, in the Cabécar auxiliary *mã* 'to go' (Margery Peña 1989: 462) and in Rama *maŋ* ~ *baŋ* 'go!', the suppletive imperative form of the verb *ta꞉k* 'to go' (Craig 1989: 169).

Future tense

In Rama, a verbal suffix *-baŋ* is used as a prospective aspect marker 'to be going to' (Craig 1989: 160). A relation of this prospective aspect marker with the verb 'go' has been explicitly argued for by Craig (ibid.: 153). Its use is illustrated in (25); I tentatively interpret it as a future tense marker here.

Rama (Craig 1989: 160)

(25) *i-tɾaːt-**baŋ***
3-walk-**FUT**
's/he is going to walk'

Likewise, the cognate Bocotá auxiliary *bɛ* indicates future tense (cf. Margery Peña 1996: 24); the Bocotá auxiliary mostly precedes the main verb, however (Margery Peña 1993: 60).

Purpose postposition

A reflex of *^{m}ba (or *^{m}bã) 'to go', Rama *baŋ*, also became a postposition indicating goal or purpose in this language (Craig 1989: 114). The purposive meaning is semantically akin to the benefactive, which is encoded by the cognate element *-mba* in Damana (cf. Trillos Amaya 2000: 754).[104]

Preverbal dative/goal marker

In Rama, an element *ba=* indicating goal or purpose is found in preverbal position. Craig (1989: 114) explicitly argues cognacy of this element with the purposive postposition *-baŋ*. The use of this element *ba=* in Rama is illustrated in (26):

Rama (Craig 1989: 79)

(26) *uːt taːɾa **ba=**i-tiŋ-i*
canoe big **GO=**3-do-IPFV
'He wants a big canoe.'

A related preverbal element <ma-> in Muisca seems to encode a generic object in (27).

[104] It remains a matter of further investigation whether or not a relation of cognacy also exists with the Kuna postposition *pa* 'with, to, from, the same, than (with comparative), by (expressing the agent)' (Holmer 1952: 112), with Boruca *maŋ* 'with, and' (Quesada Pacheco & Rojas Chaves 1999: 131, 203), and with Guaymí comitative *be* 'with' (cf. Quesada Pacheco 2008: 64).

Muisca (González de Pérez 1987: 143, adapted from Adelaar & Muysken 2004: 98)

(27) <Dios gue chie **ma**-quy-ia>
God be we **DAT**-make-PST.AGN
'God is the One who made us.'

According to the anonymous Muisca grammar from the seventeenth century, attaching <ma-> to nominalized forms of the verb 'to do, make' has the effect of "elegance" (González de Pérez 1987: 144, cited in Adelaar & Muysken 2004: 98).

Macro-Jê

In several Macro-Jê languages, it is possible to identify reflexes of a cognate root 'go', and possibly also grammaticalization paths which are parallel to those observed in Chibchan above. The vowels in the grammatical elements, however, do not seem to derive regularly from the vowels in the Proto-(Macro-)Jê verb 'to go'. Whether or not this apparent lack or regularity is due to phonetic erosion in the grammaticalized forms (cf., e.g., Heine & Kuteva 2004: 2), remains to be investigated.

A verb 'go'

Nikulin (2015b: 297) reconstructs a Proto-Jê form *mɔ̃ 'to go, come (pl.)', and Nikulin (2015a: 55) reconstructs Proto-Macro-Jê *mɔ̃ŋ 'to go, come (pl.)' (cf. also Rodrigues 1999: 200). (According to Andrey Nikulin (p.c.), novel Maxakalí data show that there was no coda consonant in the Proto-Macro-Jê form.) The reflex of the Proto-Macro-Jê form in Chiquitano is *men* 'to go' (cf. Adelaar 2008: 23).

Future tense

In Djeoromitxí (Jabutí language family), there is a future tense marker *ma* (Pires 1992: 78). In Suyá (also known as Kĩsêdjê, Northern Jê), future tense is indicated by a postposed element *mã*, as illustrated in (28).

Suyá (Santos 1997: 90)

(28) *i-tɛ̃m* ***mã***
1-go **FUT**
'I will go'

Given the formal similarity of this element *mã* with Proto-Jê 'to go', and given the frequent grammaticalization of 'go' into a future tense marker (Heine & Kuteva 2004: 161–3), it is tentatively proposed here that the Jabutí and Northern Jê future tense

markers *ma* and *mã* can ultimately be connected to Nikulin's (2015a: 55) Proto-Macro-Jê reconstruction *mɔ̃ŋ 'to go, come (pl.)' and related forms. Note, however, that Jabutí *a* and Northern Jê *ã* are not the regular reflexes of Proto-Macro-Jê *ɔ̃ (Nikulin 2015a: 42).

Allative postposition

The Suyá (Kĩsêdjê) future tense marker *mã*, illustrated in (28) is probably cognate with the Mẽbêngôkre (Northern Jê) allative case marker *mɐ̃* illustrated in (24), repeated here as (29):

Mẽbêngôkre (Reis Silva 2003: 46)

(29) *kɾĩ* ***mɐ̃*** *nẽ* *ba* *mɾuɲi* *ɔ* *tẽ*
village **ALL** FOC 1.SBJ meat make go
'I bring meat to the village.'

The Mẽbêngôkre postposition *mɐ̃* in (29) might also be interpreted as a dative or benefactive marker. In a similar vein, Apinajé *mə* has been interpreted as indicating notions of dative, benefactive, allative and locative (Oliveira 2005: 54). Ribeiro and van der Voort (2010: 536, 559) label Proto-Jê *mã, Djeoromitxí (Jabutí language family) *mã* or *ma*, and Karajá *bə̃* as dative case markers.[105] They argue that the Djeoromitxí future tense marker *ma* is likely related to the dative marker *mã* in the same language (cf. ibid.: 543, 559). Given grammaticalization paths across different languages of the world (cf. Heine & Kuteva 2007: 330), the original meaning of *mã and related forms as a case-marking postposition may have been allative.

Preverbal dative marker

The following example (30) from Mẽbêngôkre illustrates how the allative/dative postposition may come to be reassigned to preverbal position, similar to what seems to have happened with Rama *ba=* in (26) above:

Mẽbêngôkre (Reis Silva 2003: 41)

(30) *ʤɐ̃m* *a-****mɐ̃*** *pɾɐ̃m?*
INTERR 2-**DAT** want
'Are you hungry?'

[105] Elsewhere, Ribeiro (2012: 61) interprets Karajá *bə̃* as encoding a 'diffuse locative' meaning.

In a construction like (30), the dative marker *mʌ̃* may easily develop into a proclitic or prefix with respect to *prʌ̃m* 'to want'. In fact, a cognate element *mə̃-* seems to be attested in some Apinajé verbs, for instance in ***mə̃ba*** 'to fear' (cf. *aba* 'to ponder, meditate'), and in ***mə̃kɾɨ*** 'to feel cold' (cf. *akɾɨ* 'to be [physically] cold') (cf. Oliveira 2005: 134). Again, the grammaticalization chains and the developments from a postposition to a verbal prefix are parallel in certain Chibchan and Northern Jê languages. It goes without saying that these parallel grammaticalization paths are not necessarily inherited from a shared proto-language (e.g., in terms of Sapirian 'drift') but might also be the result of more recent, parallel developments.

3.3.2.3 *ⁿdaʔ ~ *taʔ and *jɛ̃ ~ *tɛ̃ 'to go$_2$'

Both in Chibchan and Macro-Jê languages, there is a second root 'to go', descending from *ⁿdaʔ ~ *taʔ in Chibchan, and *jɛ̃ ~ *tɛ̃ in Proto-Macro-Jê (compare 3 in table 114 above) (Nikulin (2015a: 55) reconstructs Proto-Macro-Jê *tɛ̃(C)). For the variation of the initial consonant in Chibchan languages, see subsection 2.4.3 above. The forms in question developed into future tense markers and into morphemes indicating allative or purposive meanings in Chibchan and in Macro-Jê languages. This is summarized in table 118.[106]

TABLE 118
POSTPOSITIONS DERIVED FROM A VERB 'TO GO' IN TWO CHIBCHAN LANGUAGES

	Térraba	Guaymí	Proto-Southern Jê
'to go'	*tɔ́* 'go, imperfective aspect' (71)	*nɔ* (107), *nigẽ* (35)	*tĩ(-g) 'to go, singular, imperfective aspect' (239)
future tense	*tɔ* ~ *tɔ́* (68)		
allative, purpose	*ɹɔ* purpose (97)	*ɾa* purpose (176)	*to allative (216)

Sources are Constenla Umaña (2007) (Térraba), Quesada Pacheco (2008) (Guaymí), and Jolkesky (2010) (Proto-Southern Jê). Page numbers in parentheses.

[106] In the context of the suppletive forms of the verb 'go' and their reflexes in Chibchan and Jê case markers, there is a certain similarity of Proto-Chibchan *ᵐba 'go$_1$' and *taʔ ~ *ⁿdaʔ 'go$_2$' with two Quechua case markers that indicate allative case or goal: *-man* (Willem Adelaar, p.c.), and *-ta*.

Chibchan

In what follows, I briefly present evidence how Proto-Chibchan *^{n}daʔ ~ *taʔ 'to go$_2$' grammaticalized as a future tense marker and as a purposive postposition in different Chibchan languages.

A verb 'to go'

Proto-Chibchan *^{n}daʔ ~ *taʔ 'to go$_2$' is reflected, for instance, in Térraba *tɔ́* 'to go, imperfective form' (Constenla Umaña 2007: 252–3), Guaymí *nã* 'to go, walk' (Quesada Pacheco 2008: 107), Pech *nã̀* 'to go' (Holt 1999a: 68), Rama *ta꞉k* 'to go' (Grinevald et al. 2002–06), Muisca <na> 'to walk' (González de Pérez 1987: 207) and <sy-> 'to go, roam' (ibid: 186).

Future tense

A reflex of Proto-Chibchan *^{n}daʔ ~ *taʔ 'to go$_2$' is found in the Térraba preverbal future tense marker *tɔ* (cf. Constenla Umaña 2007: 68), and in the Guatuso future tense marker (*futuro mediato*) |=to| (Constenla Umaña 1998: 36). The Guatuso future tense marker mostly cliticizes to the first constituent of a clause (ibid.: 181). Its use is illustrated in example (31):

Guatuso (Constenla Umaña 1998: 181)

(31) *ton=ti=**'to*** *i-kor*
I=AG=**FUT** 3.OBJ-hit
'I will hit him/her.'

Purposive postposition

As in the case of Proto-Chibchan *^{m}ba 'to go' giving rise to purposive case markers in some Chibchan languages (see above), similar phenomena seem to have occurred with the reflexes of Proto-Chibchan *^{n}daʔ ~ *taʔ 'to go$_2$': Guaymí *-ɾa ~ -ɾe ~ -dɾe* indicates purpose and follows the verb (Quesada Pacheco 2008: 175–6); it may derive from Proto-Chibchan *^{n}daʔ 'to go$_2$'. The same is true for the cognate Térraba purposive postposition *ɻɔ*, the use of which is illustrated in (32).

Térraba (Constenla Umaña 2007: 97)

(32) *ɻɔ́rɔ-ɾ* *ba* *kɔ́ŋ* *tík* ***ɻɔ***
say.PFV-1SG.SBJ 3.OBJ ALL go **PURP**
'I said it to him/her so that s/he may go.'

A Térraba preverbal element *tʰ*-, attested in ***tʰiɪ*** ‘to put’ (Sp. *meter*) (cf. *ií* ‘to put, to place’ (Sp. *poner*)) (Constenla Umaña 2007: 259, 268), may likewise be related. This implies a variation between *tʰ* and *ɻ* in Térraba. A similar variation between *tʰ* and *ɻ* occurs in Térraba *tʰéŋ* and *ɻɔŋ* both meaning ‘to be’ (Sp. *estar*), plural form (Constenla Umaña 2007: 245).[107]

Macro-Jê

In several Macro-Jê languages, it is possible to identify both forms and grammaticalization paths which may be related to those observed in Chibchan above.

A verb ‘to go’

Nikulin (2015a: 55) reconstructs Proto-Macro-Jê *tɛ̃(C) ‘to go, come (sg.)’. (Note that at present, Andrey Nikulin (p.c.) does not postulate the existence of a coda consonant in this form anymore.)

Allative/dative case marker

Proto-Macro-Jê *tɛ̃(C) ‘to go, come (sg.)’ (Nikulin 2015a: 55) seems to be related to the allative postposition *tã* of Panará, a Northern Jê language (cf. Dourado 2001: 82). The use of this Panará element is illustrated in the construction *kɾi* ***tã*** ‘to the village’ (ibid.: 168). A preverbal element *t*-, attached to the verb ‘to do, make’ may be cognate; it has been observed in some Northern Jê languages, for instance, in Apãniêkrá ***tɔ*** ‘to do, make’ (cf. Castro Alves 2004: 73), or in the causative proclitic ***tɔ***= of the same language (cf. ibid.: 76, see also above, subsection 3.3.2.1).

A cognate element may likewise be attested in Arikapú, a Jabutí language, where a preverbal element ***ʧɛ*** is analyzed as a direct-object marker by Ribeiro and van der Voort (2010: 529) in a case like (33).

Arikapú (Ribeiro & van der Voort 2010: 529)

(33)	*ahɛ=nɛ*	*paʧi*	*tə=nɛ,*	*a-nə*	***ʧɛ-iɾo***
	you=2SBJ	tobacco	bring=2SBJ	2-older.brother	**OBJ**-want

‘Did you bring tobacco? Your older brother wants some.’

This use of Arikapú ***ʧɛ***- resembles the use of Mẽbêngôkre ***-mãɾ*** before the verb ‘to want’ in example (30) above – Mẽbêngôkre ***-mãɾ*** likewise derives from a Proto-Macro-Jê verb

[107] It remains to be established whether or not Proto-Chibchan *ⁿdaʔ ~ *taʔ ‘go$_2$’ is also at the origin of the comitative case markers Térraba *tʰɔ́* (Constenla Umaña 2007: 238), Bribri [tã̀] (Chevrier 2017a: 200), and Cabécar *da ~ ɾa* (Margery Peña 1989: cvix).

'to go'. Note that Proto-Jabutí *ʧ before oral vowels regularly corresponds to Proto-Jê *j (Ribeiro & van der Voort 2010: 561), not to Proto-Jê *t.

3.3.3 Matches in verbal morphology

The following subsections discuss Chibchan–Macro-Jê parallels in the domain of verbal morphology, more specifically with regard to valency-reducing prefixes (subsection 3.3.3.1), generic object prefixes (3.3.3.2), and thematic vowel changes correlating with tense/aspect marking (3.3.3.3).

3.3.3.1 Valency-reducing *a-, *aʔ-, *a^{n}d- and *at-

Both in Chibchan and Jê languages there are reflexes of three shared allomorphs: *a-, *aʔ- and *a^{n}d- or *at- which are preverbal, valency-reducing elements. These shared allomorphs coincide in form, function, and position (compare 4 to 6 in table 114 above).

Chibchan

In what follows, I briefly present reflexes of Proto-Chibchan valency-reducing *a-, *aʔ-, and *a^{n}d- in different Chibchan languages (for more information about these elements, see subsection 2.2). All three may ultimately be related.

Proto-Chibchan *a-

Pech valency-reducing *a-* derives from Proto-Chibchan *a- and is attested in *a-pàsk* 'to leave' (compare *pàsk* 'to take out') (Holt 1999a: 60; Constenla Umaña 2012: 407). Likewise, Proto-Chibchan *a- is reflected in Kuna *a-* in a form like ***a**-pura* 'to become angry' (Holmer 1947: 117) – compare *o-pura* 'to tangle, mix'; Kuna *o-* is a causative prefix (ibid.: 117, 121), derived from Proto-Chibchan *õ 'to do, make' (see above, subsection 3.3.2.1).

Proto-Chibchan *aʔ-

Holmer (1947: 72–3) identifies a valency-reducing Kuna prefix *aʔ-* (he writes <a'->), with a final glottal stop. The glottal stop is reflected in the geminate *nn* in Kuna *annukka-* 'to wash oneself' (compare *enukka-* 'to wash [transitive]'). A prefix *aʔ- seems to be reflected in Kogi <**aḳ**veiši> 'to stir' (cf. <arveiši>, Preuss 1927: 524) and in Rama *a:burn* 'to twirl, twist, stir' (compare *alburn* 'to twirl, twist, stir') (Grinevald et al. 2002–06).

Proto-Chibchan *a^{n}d-

Proto-Chibchan *a^{n}d- is reflected in Kogi <ar->, for instance, in <**ar**veiši> 'to stir' (cf. Preuss 1927: 524). The cognate Rama prefix *al-* is a particularly prominent derivational prefix in this language – compare Rama ***al**-malŋi* 'to die' versus *malŋi* 'to kill' (Craig 1989: 74). In Muisca, a prefix <an-> encodes a passive meaning in (34).

Muisca (González de Pérez 1987: 111; adapted from Adelaar & Muysken 2004: 96)

(34) <chi-**an**-quy-squa>
1PL-**PASS**-do-IPFV
'We are being done. / It is being done to us.'

Macro-Jê

In Apinajé, a Northern Jê language, there are reflexes of three different valency-reducing prefixes *a-, *aʔ-, and *at-. Oliveira (2005: 131) identifies

- a derivational, valency-reducing prefix *a-* is attested in Apinajé ***a**gje* 'to go (pl.) in', from *gje* 'to place (pl. obj.) into deep container';
- a prefix *aʔ-* deriving intransitive ***aʔ**kapi* 'to browse' from transitive *kapi* 'to choose OBJ; select OBJ'; *aʔ-* seems to encode a generic object;
- a derivational, valency-reducing prefix *at-* is attested in ***at**katʃo* 'to tear; rip', from *katʃo* 'to tear OBJ; rip OBJ'; this prefix, too, seems to encode a generic object.

An intransitive marker *a-* is also found in Karajá and Chiquitano. The use of the Karajá intransitive prefix *a-* is illustrated in the following example (35).

Karajá (Ribeiro 2012: 88)

(35) *b-**a**-ritʃa=kɛ*
2-**INTR**-walk=POT
'Walk!'

For Chiquitano, Adelaar and Muysken (2004: 485) mention a prefix *a-* that is attached to absolute forms without an object. They explicitly notice a similarity of this intransitive prefix with Kuna *a-* mentioned above (ibid.: 64).

3.3.3.2 Generic object prefix

The Proto-Chibchan word for 'body' became a verbal prefix in Teribe/Térraba and in Kuna. It seems to have a cognate counterpart in a prefix referring to a generic object in some Macro-Jê languages (compare 7 in table 114 above).

Chibchan

Reflexes of Proto-Chibchan *apa 'body' are widely attested in the lexicon of different Chibchan languages, for instance in Bocotá *ba* 'shape, appearance' (Margery Peña 1993: 59), Bribri *apà* 'body' (Margery Peña 1982: 118), 'exterior part (of a body)' (ibid.: 125), Muisca <yba> 'body' (González de Pérez 1987: 223), and Rama *aːp* 'body' (Grinevald et al. 2002–06), among others (for more information, see the respective entry in subsection 2.2). With a more grammaticalized function, Proto-Chibchan *apa 'body' is reflected in Guatuso *ton**ɸa*** 'I myself' (Constenla Umaña 1998: 14) (compare *ton* 'I', ibid.: 13), and possibly also in the Guatuso antipassive prefix *ɸa-* ~ *ɸ-* ~ *p-* (cf. ibid.: 128–9). In Térraba, a reflex of Proto-Chibchan *apa 'body' is identified by Constenla Umaña (1981: 367) as a derivational prefix in verbs; compare the forms given in table 119.

TABLE 119
REFLEXES OF PROTO-CHIBCHAN *apa 'body' AS A DERIVATIONAL SUFFIX IN TÉRRABA

Térraba form with *pɔ́-*	Form without *pɔ́-*
pɔ́ɸ*rik* 'to tie' (232)	Térraba *ɸrík* 'to tie' (232)
pɔ́i*ɪ* 'to put on clothes' (280);	Térraba *ií* 'to put' (Sp. *colocar*) (238); *tʰiɪ* 'to put' (Sp. *meter*) (259)
pɔ́ɸ*ruk* 'to hug' (231)	Tunebo *burinro* ~ *burunro* 'to wrap, fold' (97)

Sources are Constenla Umaña (2007) (Teribe), and Headland (1997) (Tunebo). Page numbers in parentheses.

A similar phenomenon has been observed by Holmer (1947: 75) in the context of the element *ap-* in Kuna *ap-nukka-* 'to wash' (cf. *enukka-* 'to wash [transitive]'). Térraba *pɔ́-* and Kuna *ap-* are reminiscent of antipassive prefixes in that they encode a generic object.

Macro-Jê

Similar preverbal elements are also attested as valency-reducing elements in Macro-Jê languages.

In Karajá, a cognate element is probably the so-called 'antipassive' prefix described by Ribeiro (2012: 54). Its function is not to demote the patient – instead, "an unknown or irrelevant direct object" can be left unmentioned if the element *ɔ-* is prefixed to the verb (ibid.), as shown in (36).

Karajá (Ribeiro 2012: 54)

(36) *d-ãdɪ* *Ø-r-**ɔ**-θʊhɔ=r-ɛrɪ*
L-mother 3-CTFG-**GENR**-wash=CTFG-PROG
'My mother is washing (something).'

Karajá *ɔ* regularly corresponds to Proto-Jê *o (Ribeiro 2012: 260–70; Nikulin 2015a: 41), and might reflect contraction (e.g., *ap > *au > Karajá *ɔ*).

In Apinajé (Northern Jê), Oliveira (2005: 131) observes the existence of a preverbal element *ap-* in ***ap**ku* 'to eat (intr.)' versus *ku* 'to eat O[bject]', where *ap-* seems to encode an antipassive meaning or notions of 'generic thing, object'. In Apãniêkrá, valency of some transitive verbs can be reduced through a prefix *aw-* which has the allomorphs *aw- ~ ʒw- ~ a- ~ u-*. Castro Alves (2004: 77–8) interprets this Apãniêkrá prefix as referring to a generic object that is not explicitly referred to. Its use is illustrated in (37).

Apãniêkrá (Castro Alves 2004: 77)

(37) *ke* *ha* *mɛ̃* ***aw**-jahe*
3 FUT PL **GENR**-hunt
'They go hunting.'

It remains to be established to what extent the prefix in question derives from an unbound form with the meaning 'body' in Macro-Jê languages, like its presumably cognate counterpart in Chibchan languages.

3.3.3.3 Thematic vowel changes

Both in certain Chibchan and in Northern Jê languages there are cases in which similar suffixes and alternations in the thematic vowel of a verb stem are involved in tense/aspect marking (compare 8 in table 114 above).

Chibchan

In Chibchan, the alternations in question have been described in some detail for Bribri (Pacchiarotti 2013) and Cabécar (Margery Peña 1989). In both languages there is a tense/aspect which has been labelled remotospective perfective by Constenla Umaña (1981: 172). Diachronically speaking, the remotospective perfective form is the basic form of the verb stem (ibid.). Its thematic vowel is not exactly predictable in quality and it has high tone, as illustrated by some Cabécar examples in table 120. The thematic vowel of the default form of the verb stem is usually a back or low vowel, namely *ʊ* (< *a), *a*, *ã*, or *õ* (Margery Peña 1989: lxviii).

TABLE 120
CABÉCAR DEFAULT FORMS AND REMOTOSPECTIVE PERFECTIVE FORMS

Default form of the verb stem	Remotospective perfective form	English
kódʒʊ	*kódʒíɻ ~ kódʒíɻɪ*	‘to open’
kudʒʊ	*kudʒáɻ ~ kudʒáɻɪ*	‘to suck’
wʊ	*úɻ ~ úɻɪ*	‘to grind’
waka	*óɻka ~ óɻɪka*	‘to convert’
ʃʊ	*dʒíɻ ~ dʒíɻɪ*	‘to say’
tːʊ	*utéɻ ~ utéɻɪ ~ tːíɻ*	‘to speak’
kapʊ	*kapíɻ ~ kapíɻɪ*	‘to sleep’
hiʊ	*hiéɻ ~ hiéɻɪ*	‘to throw sth. in a recipient’
wítʃa	*wítíɻ ~ wítíɻɪ*	‘to make dirty’
bã	*bɛ̃́ɻ ~ bɛ̃́ɻɪ*	‘to give’
pːʊ	*búgɻíɻ ~ bukíɻ* [sic]	‘to hit with a longish object’

Source: Margery Peña (1989: lxxiii–lxxviii).

As can be seen in table 120, the correspondences between the default and the remotospective perfective forms in Cabécar are not always straightforward, for instance in a case such as *wʊ*, default form of ‘to grind’ versus *úɻ ~ úɻɪ*, remotospective perfective form. In cases such as Cabécar *tːʊ*, default form of ‘to speak’ versus *utéɻ ~ utéɻɪ*, remotospective perfective form, there are also changes on the left edge of the verb. Alternatively, these cases and others, such as *pːʊ* default form ‘to hit with a

longish object' versus *búgɽíɽ ~ bukíɽ*, remotospective perfective form, might be interpreted as root suppletion (cf. Margery Peña 1989: lxxiii–lxxviii).
The Cabécar remotospective perfective forms have an ending *-ɽ ~ ɽɪ* (Margery Peña 1989: lxxiii). The stem-final raised/fronted vowels, however, seem to be independent from final *-ɽɪ* – the same, raised/fronted thematic vowels also occur in the valency-reduced, so-called 'indefinite' forms where a suffix *-nã* /-dã/ is attached to the remotospective perfective stem: compare *ʤʊ* 'to drink' versus *ʤ**á**nã*; *ʃʊ* 'to say' versus *ʤ**í**nã*, or *kapʊ* 'to sleep' versus *kap**í**nã* (ibid.: xc).

In Kuna, Holmer (1947: 138) observes a similar alternation if the stem is followed by a suffix *-sa* (cf. also Smith 2014: 100–4): Verbal stems may end in a vowel *-a* or, less frequently, *-i*; he observes that "*a*-forms are presentic and *i*-forms preterital" (Holmer 1947: 140). This is illustrated by the forms *nakk^{w}**a*** 'to rise' and *nakk^{w}**i**-sa* 'rose', or *topa* 'to fear' versus *top**i**-sa* 'feared' (cf. ibid.: 138). The suffix *-sa* indicates the preterit (ibid.: 139); it derives from the verb *saa* 'to do, make'. Holmer (1947: 141) suggests that originally, *-a* may have encoded durative aspect, whereas *-i* indicated momentary aspect.

The evidence discussed above is probably not sufficient to reconstruct the thematic vowel alternations in question for Proto-Chibchan; further positive evidence from other Chibchan languages would be needed for this.

Jê

In Jê languages, a similar phenomenon has been observed in the context of so-called 'short forms' and 'long forms' of verbs. Example (38) shows a 'long form' in Pykobjê, a Northern Jê language:

Pykobjê (Amado 2004: 100, cited in Nikulin 2016b)

(38)	*ej-te*	*aʔjẽ*	***ko-ɾ***
	I-AG	meat	**eat-PST**
	'I ate the meat.'		

The verb *ko* 'to eat' in (38) has been combined with an element *-ɾ* which may be analyzed as a past tense marker or as a stative nominalizer, the resulting form encoding perfective aspect (cf. Ribeiro 2003; Nikulin 2016b). The default or 'short form' is used instead when aspect is not perfective, as in (39):

Pykobjê (Amado 2004: 100, cited in Nikulin 2016b)

(39) *wa* *aʔjẽ* ***ko***
I meat **eat**
'I am eating meat.'

In the nonfinite or 'long forms' of monosyllabic Apinajé verbs (to give an example from another Northern Jê language), the thematic vowel is often more fronted and raised, compared with the vowel of the short forms, as illustrated by the forms *kukja* (base) and *kukjɐɾ* (nonfinite) (Oliveira 2005: 194). Table 121 gives an overview of some cases that include a vowel change in Pykobjê.

TABLE 121
SOME 'LONG FORMS' IN PYKOBJÊ

Short or default form	Long form	English
kwa	*kwɨɾ*	'to catch'
ka	*ʧəɾ*	'to roast'
ʧwa	*ʧwɨɾ*	'to bathe'
koɾa	*koɾən*	'to kill'
aːpi	*jəpin*	'to fish'

Source: Amado (2004: 109–14).

As in some Cabécar examples above, the alternations in the Pykobjê forms in table 121 also affect the left edge of the roots, for instance in *aːpi* 'to fish' versus *j-əpi-n*. In some cases, there are also long forms ending in reflexes of *-m, *-ɲ, *-k in Northern Jê languages, but in general terms, a reflex of *-ɾ seems to be the most frequent 'long form consonant' in Northern Jê. Whether or not the endings in question may be related to the Proto-Chibchan nominalizing suffixes *-^{m}ba, *-^{n}d- and *-ka, presented in subsection 2.2, remains to be established.

3.3.4 Matches in nominal morphology

Several parallels between Chibchan and Macro-Jê languages exist in the domain of nominal morphology. These cases will be discussed in the following subsections. First, I present and discuss parallel sets of case-marking postpositions indicating, for instance,

locative, instrumental, and dative case. This is followed by a description of parallels in derivational elements. The subsection closes with a presentation and discussion of Chibchan–(Macro-)Jê sets of shared possessive markers and sets of deictic elements.

3.3.4.1 Postpositions

Chibchan and Macro-Jê languages share a large set of case-marking postpositions which will be discussed in the present subsection: locative *ki; allative/dative *ka; locative *da ~ *ta (Chibchan) or *jɛ ~ *tɛ (Macro-Jê); and instrumental/locative *^{n}di (Chibchan) or *ji ~ *ɾi (Macro-Jê) (compare 9 to 12 in table 114 above).

Locative: *ki

Both Chibchan and Macro-Jê languages have a locative postposition in common that derives from *ki (compare 9 in table 114 above).

Chibchan

The reflex of a Proto-Chibchan locative marker *ki is attested in several Chibchan languages: There is Rama *ki* 'in', which occurs as a case-marking postposition and as a relational preverb (Craig 1989: 114), and Kuna *ki*, a postposition with the meaning 'in, with, at, about, like' (Holmer 1947: 186; 1952: 52). The use of Kuna *ki* is illustrated by Holmer (1947: 186) with expressions such as *uluk:a* ***ki*** 'in the body', *uarsi* ***ki*** 'with an arrow', *pe* ***ki*** 'like you', and *e* ***ki*** [egi] 'in it' (the latter form is compared, by Holmer (ibid.) with Kogi *(h)ai-ki* 'this way'). Proto-Chibchan *ki locative might be etymologically linked with another Proto-Chibchan postposition: *kĩ 'for, because', which is only tentatively reconstructed here. The latter has reflexes in several Chibchan languages too, all bearing traces of the original vowel nasality: Rama *kiŋ* benefactive (both as a case-marking postposition and as a relational preverb) (Craig 1989: 114), Térraba *kʰiŋ* 'for, because' (Constenla Umaña 2007: 268), and Cabécar *kĩ́* with a meaning 'in', 'on', 'to' and 'because' (Margery Peña 1989: 147).

Macro-Jê

Xokléng and Kaingáng (both Southern Jê) have a locative marker *ki* 'in, inside, on' (Henry 1948: 202; Wiesemann 1972: 207), and in Karajá there is a locative postposition *kɪ* (Ribeiro 2012: 24).[108] Cognacy of Southern Jê *ki* with Karajá *kɪ* was proposed both by

[108] The Karajá locative postposition *kɪ* is labelled 'stationary locative' by Ribeiro (2012: 61). It is particular, given that *k* in female speech regularly corresponds to zero in male speech (see also 3.3.1.2 above). Notwithstanding, the locative postposition *kɪ* remains unaffected by this rule (Ribeiro 2012: 24).

Rodrigues (1999: 200) and Ribeiro (2012: 273). The use of Karajá *kɪ* is illustrated in (40):

Karajá (male speech) (Ribeiro 2012: 226)

(40) *ɗəkɪ* *θohodʒi* *ɨwa***=kɪ** *Ø-d-a-ɾɪ=d-e*
he one *jatobá***=LOC** 3-CTPT-INTR-leave=CTPT-IPFV
'He was left alone on the *jatobá* tree.'

Allative/dative *ka

Both Chibchan and Macro-Jê languages share dative or allative markers that derive from a proto-form *ka (compare 10 in table 114 above).

Chibchan

The reflexes of Proto-Chibchan *ka allative/dative often imply movement rather than stative location and are found in several daughter languages. Among the reflexes of the case-marking morpheme in question are Bocotá *kɛ* ~ *gɛ* dative postposition (Margery Peña 1996: 19), Guatuso *=ko* 'in, inside, to, toward (locative, ilocative, alocative, temporal)' (Constenla Umaña 1998: 107), Kuna *ka* dative ('to, for, against') (Holmer 1947: 186), and Muisca <c(a)> locative (goal) (Ostler 1993). The use of Guatuso *=ko* is illustrated in (41):

Guatuso (Constenla Umaña 1998: 107)

(41) *i-ti'o=ŋe* *'u***=ko**
3.SBJ-enter-REA house**=LOC**
'S/he entered the house.'

A cognate element in Ika is the proclitic *kə=*. Frank (1985: 93) interprets it as a valency-increasing verbal element indicating a peripheral participant. Its use is illustrated in (42) where it is attached to a verb and follows the locative marker *-seʔ*.

Ika (Frank 1985: 94)

(42) *kafe* *Pablo-seʔ* ***k*****=***əŋeiʔ-na-ɾua* *ni.*
Coffee Pablo-LOC **PP=**sell-PTCP-1SG.SBJ AFF
'I sold coffee to Pablo.'

In fact, this Ika element indicating a peripheral participant, according to Frank's (1985) interpretation, seems to derive from the allative/dative postposition *ka which was reassigned to preverbal position.

Jê

An allative/dative case marker *ka*, which was reassigned to preverbal position is probably also attested in the so-called *ka*-verb stems in Apinajé. They strikingly resemble the Ika case in (42) above. Compare, for instance, Apinajé *o* 'to suck' versus ***ka****ʔo* 'to suck out of fruit', *prepɾek* 'fast, quick; to rush' versus ***ka****prepɾek* 'to spank' (Oliveira 2005: 118, 406), Proto-Jê *wẽ 'to speak' versus Apinajé ***ka****pẽ* 'to talk to' (Oliveira 2005: 388; Ribeiro & van der Voort 2010: 552). In Apinajé, the transitive verbs that contain the element *ka-* imply physical contact or manipulation (Oliveira 2005: 117). In Proto-Northern Jê, one may compare *jakô 'to blow' versus ***ka**kô 'to blow into', or *põ 'to rub' versus ***ka**põ 'to sweep'. Whether or not there is a relation of cognacy with the element *kã* in Kaingáng ***kã****tĩŋ* 'to reach, come' remains to be established (the right-hand element of the Kaingáng form is cognate with Apinajé *tẽm*, Proto-Jê *tẽm 'to come', long forms) (cf. Ribeiro and van der Voort 2010: 552).

Locative: *^{n}da ~ *ta and *jɛ ~ *tɛ

Chibchan and Macro-Jê languages share postpositions that derive from a proto-form *^{n}da ~ *ta and *jɛ ~ *tɛ, respectively (compare 11 in table 114 above).

Chibchan

In different Chibchan languages, there are reflexes of a locative case marker *^{n}da or *ta. The initial consonant is difficult to determine in some cases: Guatuso *ta* 'in, to (temporal)', for instance, may derive from either of both proto-forms. The use of the Guatuso postposition is illustrated in (43):

Guatuso (Constenla Umaña 1998: 113)

(43) *na-ˈluɾi=to* *ˈtoxi* *Ø-tuˈe* ***ta***
1.SBJ-come back=FUT sun 3.SBJ-go **in**
'I will come back in the afternoon (lit. "when the sun is leaving")).'

Reflexes of Proto-Chibchan *^{n}da are Muisca <n(a)> location (and rest in it) (Adelaar & Muysken 2004: 99), the Cabécar postposition *nã* /dã/ 'in' (cf. Margery Peña 1989: 200), and Kogi *-la* in *ˈhula* 'in the house' (Ortiz Ricaurte 2000: 770). Instead of *^{n}da, a reflex

of Proto-Chibchan *ta is found in Boruca *ta* 'in' (Quesada Pacheco & Rojas Chaves 1999: 140).

The reflex of the Proto-Chibchan locative marker *ta probably developed further into an agent or ergative marker in some Chibchan languages. The development of a locative marker into an element indicating the agent is attested across several languages of the world (Heine and Kuteva 2004: 199–200). In Chibchan the cases in question are the Bribri ergative postposition *tȍ* (cf. Constenla Umaña 2012: 407) and probably also Tunebo *-at* ergative (Headland 1997: 14). (That the Tunebo form *-at* derives from *ta by metathesis is conceivable given the observation that the reflex of Proto-Chibchan *ka allative/dative seems to be Tunebo *ak* 'for, to', cf. Headland (1997: 62).)

Macro-Jê

Proto-Chibchan *^{n}da ~ *ta locative seems to have a cognate counterpart in Proto-Macro-Jê which is tentatively reconstructed here as *jɛ ~ *tɛ. The first of these elements seems to be reflected in Proto-Jabutí *ʧɛ locative (cf. Ribeiro & van der Voort 2010: 543). The Djeromitxí reflex of Proto-Jabutí *ʧɛ is illustrated in (44):

Djeoromitxí (Ribeiro & van der Voort 2010: 536)

(44) *nipɛru* ***ʧɛ*** *a-runã*
bench **LOC** 2-sit
'You sit on the bench.'

Proto-Jabutí *ʧ (before oral vowels) regularly correspond with Proto-Jê *j (Ribeiro & van der Voort 2010: 561) and probably derives from Proto-Macro-Jê *j.

As to Proto-Macro-Jê *tɛ, some Macro-Jê languages have an ergative marker *tɛ* which may eventually turn out to be related: compare, for instance, Apinajé *tɛ* (Oliveira 2005: 178), Timbira *tɛ* (Castro Alves 2004: 86), Pykobjê *te* (Amado 2004: 36), Krahô *tɛ* (Souza 1990, cited in Amado 2004: 58), and, provided that the Macro-Jê status of this language is accepted, Maxakalí ergative *-te* [ˈtɛʔ] (Campos 2009: 92); in this case, one might tentatively reconstruct a Proto-Macro-Jê ergative marker or agent disambiguator *tɛ.

Instrumental/locative *^{n}di and *ji ~ *ɾi

An instrumental/locative postposition *^{n}di and *ɾi ~ *ji can be reconstructed for Proto-Chibchan and for Macro-Jê, respectively (compare 12 in table 114 above).

Chibchan

Proto-Chibchan *^{n}di is reflected, for instance, in Kogi *-li* locative (Ortiz Ricaurte 2000: 772) and in the Kuna instrumental/locative postposition *ti* (Holmer 1947: 63). The cognate Chimila suffix *-ɾi* can be used to disambiguate the agent in a transitive construction and has been interpreted as a ergative–dative marker (Adelaar & Muysken 2004: 79). Likewise, cognate Guatuso *ti* indicates ergative case in certain transitive constructions, such as (45):

Guatuso (Constenla Umaña 1998: 115)

(45) *oˈʧ̂ipaˈka* ***ti*** *mi-kuˈa=ŋe*
man **AG** 2.OBJ-see=REA
'The man saw you.'

With respect to Bribri, Pacchiarotti (2016) mentions the existence of *di*, an ergative postposition only used with third-person referents in story-telling. Whether or not Teribe *dí* also belongs here needs further investigation. This element indicates the subject in inverse constructions (cf. Quesada 2000a: 63). Its use is illustrated in the following example (46):

Teribe (Quesada 2000a: 12)

(46) *ta* *ĩ-ja* *Juan* ***dɪ***
1SG see-IPFV.INV Juan **SBJ**
'Juan sees me.'

Macro-Jê

In Macro-Jê languages there are instrumental/locative markers derived from *ji or *ɾi. In Karajá, the clitic *=dɪ* has an instrumental meaning. Its use is illustrated in (47):

Karajá (Ribeiro 2012: 102)

(47) *hãwəkɔ=**dɪ***
canoe=**INSTR**
'with the canoe'

Karajá *d* regularly corresponds to Proto-Jê *j (Ribeiro 2012: 271), and may derive from Proto-Macro-Jê *j. Instead, Apãniêkrá, a Northern Jê language, has a locative marker *-ɾi*, illustrated in example (48):

Apãniêkrá (Castro Alves 2004: 87)
(48) *ku-ɾi*
3- **LOC**
'in that place'

It is tentatively proposed here that Apãniêkrá *ɾ* derives from Proto-(Macro-)Jê *ɾ. Whether or not the Proto-Jabutí dative postposition *ri (Ribeiro & van der Voort 2010: 543) belongs here as well remains to be established.

3.3.4.2 Elements relating to shape

A peculiar feature of Chibchan languages are elements derived from *k^{w}a 'seed, fruit' and *ka^{n}d- ~ *kat- 'stick, bone, tree'. Elements derived from these terms can have several functions in Chibchan languages, for instance as numeral classifiers relating to shape, or as noun classifiers or class terms (e.g., Constenla Umaña 1988; Pache 2016a). These Proto-Chibchan elements seem to have cognate counterparts in (Macro-)Jê languages, as will be illustrated in the following two subsections (compare set 13 in table 114 above). Notwithstanding, they do not seem to have any reflexes in all Macro-Jê languages (e.g., there is no reflex of them in Chiquitano), nor in all Chibchan languages, either. The exact origin of the parallels described below is yet to be determined.

*k^{w}a 'three-dimensional entity'

Classifying or derivational elements derived from a proto-form *k^{w}a may relate to roundish entities in Chibchan and in Macro-Jê languages.

Chibchan

An element derived from *k^{w}a is attested as a noun classifier or class term in most Chibchan languages (cf. Pache 2016a: 448), with the exception of Rama and Guatuso. The meaning of the unbound Proto-Chibchan element *k^{w}a is reconstructed here as 'seed, fruit'. In several cases, such as Kuna *ak**k^{w}a*** 'stone' (ibid.), or Muisca <u**pqua**> 'eyes' (González de Pérez 1987: 285), the reflex of *k^{w}a is attached to an etymologically transparent root (in these specific cases, reflexes of Proto-Chibchan *hak ~ *kaʔ 'stone' and *u^{m}ba 'face, eye, fruit', cf. subsection 2.2). In Proto-Chibchan, a derivational suffix *k^{w}a is attested in *siʔ-k^{w}a 'pale', derived from *siʔ 'moon', and in [*^{n}di-g^{w}a], probably /*^{n}di-k^{w}a/ 'lake, rain', derived from *^{n}diʔ 'water, river'. In these cases, it is a derivational suffix rather than an element encoding shape. An example of a

numeral classifier derived from *k^{w}a ‘seed, fruit’ is Guaymí *kwɔ- ~ gwɔ-*, which is prefixed to number terms referring to round objects and years (Quesada Pacheco 2008: 91). In Colombian Chibchan languages there are some traces of fossilized classifiers likewise derived from Proto-Chibchan *k^{w}a, for instance in Muisca <qhu**pqua**> ‘seven’ (cf. González de Pérez 1987: 162). More information and examples of *-k^{w}a* and related forms as a noun classifiers or numeral classifier in Chibchan languages can be found in the literature (e.g., Constenla Umaña 1988; Pache 2016a).

Macro-Jê

A cognate element seems to exist in Macro-Jê languages: *ko in Proto-Jabutí *ʧa**ko** ‘mouth’, *kua in Proto-Jê *j-ar**kua** ‘mouth’ (cf. ibid.: 557–8). The status of *-ko and *-kua as suffixes is corroborated by a comparison of Chiquitano <arù-s> ‘lip’ (ibid.: 361) with Proto-Jê *j-**ar**kua ‘mouth’ (cf. Adelaar 2008; Ribeiro & van der Voort 2010). The Chiquitano form does not contain any corresponding element reflecting *kua. In the Proto-Jê and Proto-Jabutí forms for ‘mouth’, the reflex of *kua may refer to round shape irrespectively of convexity or concavity.

In Panará, a northern Jê language, Dourado (2001: 208) mentions the classifiers *kua* and *kwə* for fruit, and *krɛ* for concave objects which are probably all etymologically related.[109] The word *kua* ‘fruit’ can be used in this language as a classifier which is incorporated in the verb form as an element which is coreferential with the subject, as illustrated in (49):

Panará (Dourado 2001: 24)

(49) <nãkowsia ji**=kua**=tẽ>
orange.SBJ REA.INTR**=fruit=**fall
‘The orange fell down.’

***ka^{n}d ~ *kra ‘one-dimensional entity’**

Classifying or derivational elements derived from a proto-form *ka^{n}d or *kra frequently relate to longish, stick-like entities in Chibchan. Cognate elements may also exist in Macro-Jê languages, although apparently not as suffixed derivational elements.

[109] In some instances, there seems to be an alternation or variation between *k^{w}V ~ kuV ~ krV*-sequences within Jê: final *kua, which may also have had as a meaning ‘hole’ in Proto-Jê *j-ar**kua** ‘mouth’, may be compared with Proto-Cerrado Jê (Proto-Amazonian Jê) *krɛ in *ʃĩya-**krɛ** / *ɲĩya-**krɛ** ‘nose’ (cf. Nikulin 2015a: 71). Whether or not a similar alternation or variation *g^{w}V ~ guV ~ grV* also exists in Chibchan remains to be established; a possible case is Muisca <gue> [we] ‘twenty’ (Adelaar & Muysken 2004: 106), which may eventually turn out to be related to Guaymí *gre* ‘twenty’ (cf. Quesada Pacheco 2008: 92).

Chibchan

In Chibchan languages, there is a classifying element derived from *-ka^{n}d- (ultimately derived from *ka^{n}d ~ *kat- ‘stick, bone, tree’). In the Proto-Chibchan forms reconstructed here, this element is attested in Proto-Chibchan *^{m}bu^{n}d-ka^{n}d(a) ‘wind’, derived from *^{m}bu^{n}d- ‘breath, wind’. Some examples of its reflexes as a noun classifier or class term are Rama *k^{w}iːka-**kat*** ‘arm’ (*k^{w}iːka* ‘hand’) and *k^{w}siŋ-**kat*** ‘ax handle’ (Craig 1989: 65; Grinevald 2000: 60) or Térraba ***maŋkɾá*** ‘knife’ (Constenla Umaña 2007: 133) (compare Cabécar *bak* ‘axe’ (Margery Peña 1989: 450) which is probably cognate with the left-hand element of the Térraba form). Elements derived from *-ka^{n}d are relatively rare among the fossilized numeral classifiers of Colombian Chibchan languages; among the few exceptions is Ika *mai**kəni*** ‘three’ (cf. Huber & Reed 1992: 181).

Reflexes of *ka^{n}d- ~ *kat- ‘stick, bone, tree’ also became instrumental nominalizers and/or markers of purpose in some Chibchan languages, such as Cabécar, Guaymí, Kuna, or Teribe/Térraba. Two grammaticalization chains are conceivable: ‘stick’ > ‘thing’/‘belonging’ > possessive marker > benefactive case marker > purpose marker > instrumental nominalizer and ‘stick’ > ‘thing’/instrumental nominalizer > purpose marker (Pache 2016a: 442–7). An example of a purpose marker derived from *ka^{n}d- ~ *kat- ‘stick, bone, tree’ in Kuna is (50).

Kuna (Holmer 1947: 187)

(50) *sikke–**kala***
cut–**PURP**
‘in order to cut’

For further examples and a detailed discussion of grammaticalization paths see Pache (2016a).

Jê

In some Northern Jê languages, an element *kra* or related form is attested, possibly with a meaning ‘stick-like entity’, as a left-hand element in ‘axe’ terms: Rodrigues (1999: 199) gives Northern Jê ***kɾɐ̃***-*mɛ̃n* ‘axe’, Krenak ***kra***-*pok* ‘axe’ and Purí ***kra***-*maŋ* ‘axe’ (segmentations and transcriptions by Rodrigues, ibid.). The coincidence of the right-hand element in these forms with Southern Jê *mɛŋ* ‘axe’, observed by Rodrigues (1999: 199) is remarkable. It goes without saying that ‘axe’ terms are not part of basic vocabulary. Therefore, the Chibchan forms reflecting *ka^{n}d- ~ *kat- ‘stick, bone, tree’ and left-hand elements in these Northern Jê ‘axe’ terms are not necessarily inherited from a common ancestor language shared by Chibchan and (Macro-)Jê languages.

Additionally, there is also some similarity between purpose markers which are derived from *ka^{n}d in some Central American Chibchan languages (see above), and future tense markers in some Macro-Jê languages: compare the Karajá future tense marker =*kəɾe* (cf. Ribeiro 2012: 58) (not subject to the *k*-dropping rule in Karajá male speech, cf. ibid.: 143) and the Krahô (Northern Jê) future tense marker -*kɾa* (Popjes & Popjes 1986: 179). Whether or not this coincidence with Chibchan languages is fortuitous remains to be established.

3.3.4.3 Relational elements

Formal parallels between Macro-Jê and Chibchan languages also exist in the domain of relational elements that may be (fossilized) third-person markers, or that have a linking function between the constituents of possessive constructions. The elements in question, shared by Chibchan and Macro-Jê languages, are *^{n}d and *ʦ and related forms (compare 14 and 15 in table 114 above). In Macro-Jê languages, these morphemes have been addressed as 'relational prefixes' by Rodrigues (2000, 2009).

Chibchan

In Chibchan languages, two different 'relational' elements can be reconstructed, *^{n}d(i)- and *ʦ(i)-. In modern Chibchan languages, these elements are reflected in fossilized affixes (possibly third-person markers, originally) and in elements inserted between the head and the dependent of possessive constructions.

Proto-Chibchan relational$_1$ *^{n}d(i)- is reflected, for instance, in Muisca <**n**iomy> 'testicle' versus <iomza> 'potato' (González de Pérez 1987: 331), Damana ***ni**kuma* 'egg' (Huber & Reed 1992: 260) versus Boruca *kúp* 'egg, testicle' (Quesada Pacheco & Rojas Chaves 1999: 153, 193), Boruca *ɾunkáx* versus *unká?* 'father-in-law' (Quesada Pacheco & Rojas Chaves 1999: 191). A reflex of relational *^{n}d(i) may also exist, in fossilized form, in Proto-Chibchan *u^{m}ba ~ *^{n}**d**u^{m}ba 'brother-in-law', *^{n}**d**u 'father, uncle, ancestor' (compare Proto-Jê *j-um 'father' (Ribeiro & van der Voort 2010: 558)), *^{n}**d**ii(k) 'nose' (compare Proto-Jê *j-ĩja 'nose' (Ribeiro & van der Voort 2010: 569)), and *^{n}**d**u? 'tooth' (compare Proto-Jê *j-ua 'tooth' (Ribeiro & van der Voort 2010: 558)). In some few cases, a reflex of *^{n}d(i) occurs between necessarily contiguous morphemes, for instance in Atanques <uma**n**júma> 'eyelid' (Celedón 1892a: 597) (cf. <úma> 'eye' (ibid.: 597) and Proto-Chibchan *huBa 'skin'). Also, in Damana, Trillos Amaya (1999: 26–8) observes the existence of a suffix -*n* that she interprets as a connective element. Its use, connecting a direct object with a following, nominalized verb is illustrated in (51).

Damana (adapted from Trillos Amaya 1999: 27)
(51) *aiju-**n*** *gok-ka*
coca-**L** make-PTCP
'coca harvesting'[110]

If followed by a voiceless stop, this Damana linking element has a phonologically conditioned allomorph *-h* (ibid.: 27–8), as shown in (52), where it follows a person-marking element:

Damana (adapted from Trillos Amaya 1999: 28)
(52) *mə-**h**-tekoʃ-kua*
2-**L**-prepare.the.soil.for.sowing-2PL.PTCP
'You prepare the soil for sowing.'

Proto-Chibchan relational$_2$ *ʦ(i)- is a linking element. I propose that it is only reflected between contiguous elements forming a compound or possessive construction, such as Rama *ups**i**ri* 'tears' ('water (*ri*) of the eye (*up*)') (cf. Craig 1989: 66), and in Cabécar *huɻa**s*** *kˀó* 'finger' ('leaf (*kˀó*) of the hand (*huɻa*)') (cf. Margery Peña 1989: 419). The reflex of the Proto-Chibchan linking morpheme *ʦ(i)- is <s> or <z> in Muisca. A reflex <s> is found in Muisca <cha-**s** gue> 'the man's house' (González de Pérez 1987: 137, cited in Ostler 1994: 208; <cha> 'man, male', <gue> 'house').

In several cases, however, no element is inserted between the head and the dependent of a possessive construction, for instance in Kuna *wa-kala* 'face, cheek' (Holmer 1952: 180) (= 'bone of the face'), and in Chimila /bittakɾa/ [^{m}bít·ākɾà] 'chest' (Malone 2005: 202) (= 'bone of the heart', cf. Proto-Chibchan *^{m}bih^{n}da ~ *^{m}bihta 'heart, liver, center'). I have no explanation yet for this phenomenon.

Likewise, it remains to be established to what extent a relational element or genitive marker *i can be reconstructed for Proto-Chibchan. Such an element might have a reflex in final *i* in Kogi possessive pronouns such as *na'hi* 'my' (compare *nas* 'I') (Ortiz Ricaurte 2000: 771).

Macro-Jê

In Macro-Jê languages, elements that have been labelled 'relational prefixes' are widely attested (e.g., Rodrigues 2000, 2009; Ribeiro 2004). An element that seems to be cognate with Proto-Chibchan *^{n}d(i)- is Proto-Jê *j-, attested in Proto-Jê *j-ar- 'mouth'

[110] The interpretation of *-ka* as indicating a present participle is based on Constenla Umaña (2012: 406).

and *j-õtɔ ‘tongue’, which is cognate with Karajá *d-* (Ribeiro & van der Voort 2010: 558–9). This element is used in contiguous constructions such as (53):

Karajá (Ribeiro & van der Voort 2010: 551)
(53) *dɔɾɛ* ***d****-e*
parrot **L**-wing
‘the parrot’s wing’

For Proto-Jabutí, Ribeiro and van der Voort (2010: 565–6) reconstruct two relational elements (linking prefixes) **ʧ-* and **n-* that both regularly correspond to Proto-Jê **j-* (followed by an oral or nasal vowel, respectively). The cognate forms in Apinajé (Northern Jê) are given below. Example (54) illustrates the linking element ***ʧ****-*:

Apinajé (Ribeiro 2011: 113)
(54) *it-****ʧ****-wa*
1-**L**-tooth
‘my tooth’

(55) is an example illustrating the use of the linking prefix ***ɲ****-* in Apinajé:

Apinajé (Oliveira 2005: 183)
(55) *i-****ɲ****-õʔtɔ*
1-**L**-tongue
‘my tongue’

The corresponding elements in Ofayé are reflected in the forms ***ʃ****-ɛr* ‘mouth’, and ***j****-õra* [ɲõˈra] ‘tongue’ (Ribeiro & van der Voort 2010: 566). In Chiquitano, two similar elements *s* and *ɲ* occur only in the context of the first person, for instance in Chiquitano (Besɨro) *nisaru* ‘my mouth’ (word-initial *n-* is a euphonic element) and *niɲutu* ‘my tongue’ (Galeote Tormo 1993: 88). They have been argued to be cognate with the Ofayé and Apinajé elements mentioned above by Adelaar (2008: 24) and Ribeiro (2011: 113), respectively, but not by Nikulin (forthc.).

Finally, in Djeoromitxí (Jabutí language family), a variation between initial *ɾ* and *h* (Djeoromitxí *ɾ* and *h* regularly correspond with Proto-Jê **j*) has been observed in forms like *ɾapa* ~ *hapa* ‘arm’, *ɾu* ~ *hu* ‘father’, or *ɾʉ* ~ *hʉ* ‘tooth’ (Pires 1992: 45;

Ribeiro & van der Voort 2010: 565).[111] An alternation of *n* or *ɾ* with *h* has also been observed in certain Djeoromitxí verbs (*nõkʉ* ~ *hõkʉ* 'to fall', *ɾʉmi* ~ *hʉmi* 'to be ill/hurt', *ɾabə* ~ *habə* 'to become tired' (Sp. *cansar*) (cf. Pires 1992: 46; Ribeiro & van der Voort 2010: 532–3). This alternation between the linking morpheme *ɾ* or *n* and *h* depends on the element that precedes the form in question (a prefix indicating person, or a left-hand element in a compound entails the use of *ɾ* or *n*) (van der Voort 2007: 142). This is reminiscent of the alternation that has been described above for Damana. Whether or not this parallel is entirely due to later, independent developments in the languages, remains an open question.

In non-contiguous forms, a relational prefix or third-person marker *s- (occurring with vowel-initial stems) ~ *i- (occurring with consonant-initial stems) was used in Proto-Jê. These two allomorphs are reconstructed as such by Ribeiro and van der Voort (2010: 551); the counterpart of Proto-Jê relational *s- in Karajá is *ɗ-* (ibid.). Its use is illustrated in (56).

Karajá (Ribeiro & van der Voort 2010: 551)

(56) ***ɗ-e***
3-wing
'its wing'

All these relational elements in Macro-Jê languages have cognate counterparts in Tupían and Cariban, according to Rodrigues (2000, 2009). The forms described here for Proto-Macro-Jê are interpreted as a case of shared allomorphy by Rodrigues (ibid.). Based on this and other parallels with Tupían and Cariban, this author proposed the existence of a genealogical relationship between Tupían, Cariban and Macro-Jê languages. Four different relational prefixes were reconstructed by Rodrigues (2009: 165) for Proto-Tu-Ka-Jê, as shown in table 122.[112]

[111] Comparing the two last Djeoromitxí forms with their counterparts in Proto-Chibchan, namely *^{n}du 'father, uncle, ancestor' and *^{n}duʔ 'tooth', suggests that initial *h* is secondary in the Djeoromitxí forms in question, if *ɾ* and *h* are allomorphs.

[112] Viegas Barros (2005b) detects similar allomorphs of relational elements in Guaicuruan languages.

TABLE 122
TU-KA-JÊ RELATIONAL PREFIXES

	Vowel-initial	Consonant-initial
Contiguity	*D- (a coronal stop)	*Ø-
Non-contiguity	*ʦ-	*i-

Source: Rodrigues (2009: 165).

Whereas a reflex of the coronal stop is attested in contiguous forms in Macro-Jê languages, it is predominant in Chibchan non-contiguous forms, however. Instead, a reflex of the affricate is attested in non-contiguous forms in Macro-Jê languages, but in contiguous forms in Chibchan languages.

3.3.4.4 Deictic elements

Both in Chibchan and in Macro-Jê languages, a deictic element *iC seems to refer to proximate entities, whereas an element *a is used to refer to more distant entities (compare 16 and 17 in table 114 above).

Chibchan languages

For Chibchan languages, Constenla Umaña (2012: 408) reconstructs two forms for the third person marker: *hi and *ha, the first being much more frequent than the second. The demonstrative element / third person marker *hi ~ *iʔ which is reconstructed in the present work (see subsection 2.2) has reflexes in several Chibchan languages, shown in table 123.

TABLE 123
SOME REFLEXES OF PROTO-CHIBCHAN *hi ~ *iʔ ‘this’

Language	Meaning
Boruca	*i* ~ *iʔ* ‘he, she, it’ (Quesada Pacheco & Rojas Chaves 1999: 140), *í* ‘what, which’ (ibid.: 133; 181)
Bribri	*iʔ* ‘this’ (Margery Peña 1982: 124), *ì* ‘what?’ (ibid.: 147)

Cabécar	*hí* 'this' (Margery Peña 1989: 104)
Kogi	*hi* 'what?' (Ortiz Ricaurte 2000: 771)
Kuna	*i* 'this' (cf. Holmer 1947: 106)
Pech	*ĩ* 'this, these (proximate)' (Holt 1999a: 62)

An element *a (my own reconstruction) has reflexes in the following Chibchan languages, shown in table 124.

TABLE 124
SOME REFLEXES OF PROTO-CHIBCHAN *a 'that'

Language	Meaning
Barí	[ʔā] 'he, she' (cf. Huber & Reed 1992: 197)
Damana	*a-* third person object (Trillos Amaya 2000: 755)
Duit	<a-> 'he, his' (Lehmann 1920: 54)
Guatuso	*o꞉* relative pronoun (Constenla Umaña 2012: 408)
Kuna	*a* 'that' (cf. Holmer 1947: 106)
Muisca	<a-> third person marker (Adelaar & Muysken 2004: 97)
Pech	*ã* 'that, those (distal)' (Holt 1999a: 62)

In Kuna and Pech, the form derived from *hiʔ ~ *iʔ is clearly associated with the notion of proximity, whereas the form derived from *a is connected with the notion of spatial distance: compare Kuna ***i****tti tule* 'this man' versus ***a****ti ome* 'that woman' (*tule* 'man', *ome* 'woman', cf. Holmer 1946: 190).[113] Holmer (1947: 106) interprets final *-til-tti* as a nominalizing element or a "general noun-forming suffix" (ibid.: 167).

Northern Jê

In Jê languages, a parallel association of *i* proximate demonstrative and *a* distal demonstrative is found in Apãniêkrá (Northern Jê) (57).

Apãniêkrá (Castro Alves 2004: 78)

(57) ***i****-ta* *ri*
1-DEM LOC
'here, at this place'

[113] As a matter of fact, Greenberg (1987: 286) explicitly proposes an element *a* as a "third-person pronoun and distance demonstrative" for his alleged Chibchan–Paezan subgroup of Amerind.

Castro Alves (2004: 78) interprets this demonstrative construction as consisting of a person-marking element (*i-* first person) which is prefixed to a demonstrative root. In contrast, the distal demonstrative is built on a demonstrative root to which a second-person prefix *a-* is attached, as illustrated in the following example (58).

Apãniêkrá (Castro Alves 2004: 79)
(58) ***a****-ta* *ɾi*
2-DEM LOC
'over there, in that place'

In the two Apãniêkrá constructions shown above, Castro Alves (2004) interprets *i* and *a* as person-marking prefixes, namely *i-* first person and *a-* second person. Likewise, in Apinajé, first person *ic-*/*iɲ-* and second person *a-* can be attached to a deictic element *taɾ*, in order to express the notions 'here' and 'there', according to the interpretation of Oliveira (2005: 164). Ribeiro and van der Voort (2010: 551) reconstruct a Proto-Jê first-person marker *ij-. Jolkesky (2010: 249) reconstructs first person *iɻ for Proto-Southern Jê. For Pre-Chiquitano, Nikulin (forthc.) reconstructs two first-person singular markers, *iʂ- feminine, and *ij- masculine. Whereas in the modern northern Jê languages Apãniêkrá and Apinajé, deictic elements have been interpreted as derived from prefixes indicating the first and second person, the inverse process may have taken place originally, namely Proto-Macro-Jê verbal person marking prefixes (*ij- 'first person', *a- 'second person') being derived from deictic elements.

The oposition between *a* and the high front vowel *i* might be a frequent way in which languages symbolize real-world oppositions. Notwithstanding, Chibchan and Macro-Jê coincide in that the *i*-form refers to the deictic center, whereas the *a*-form refers to more distant entities.

3.3.5 Regular sound correspondences

In the establishment of a genealogical relationship between two languages, a sufficient number of recurrent sound correspondences occurring in basic morphology and vocabulary is an important argument (for a discussion of this requirement, see subsection 2.1 above). In terms of an overview, table 125 presents some potential Chibchan–Macro-Jê cognate sets, including those attested in the morphological parallels of the previous subsections.

TABLE 125
(PROTO-)CHIBCHAN–(PROTO-)(MACRO-)JÊ COGNATE SETS REFLECTING RECURRENT SOUND CORRESPONDENCES

Set	Proto-Chibchan	Proto-Macro-Jê	Proto-Jê, Proto-Cerrado Jê, Proto-Northern Jê	Single languages	Function/meaning
1	*a-, *aʔ-, *a^{n}d-	*a-		Apinajé *a-*, *aʔ-*, *at-* (Ol 131), Karajá *a-* (R 88)	valency-reducer
2	*a	*a	*a- PJ (RV 551)		PC distal demonstrative; PMJ, PJ 2nd person
3	*ka		*ka- PNJ, attested in several verbs		PC allative/dative; PJ indirect-object prefix
4	*ka	*pa(C) (Na 62)			PC ‘leaf’; PMJ ‘arm, branch’
5	*ka^{n}d- ~ *kat-	*par$^{\circ}$ (Na 60)			PC ‘stick, bone, tree’; PMJ ‘foot’; cf. also PNJ *krat ‘base, trunk’ (NS), *kra- and related forms in NJ, Krenák and Purí ‘axe’ terms (Roa: 199)

6	*-k^{w}a	*-k^{w}a	*-kua PJ (cf. RV 558), possibly also PCJ *-kɾɛ (cf. Na 71)	PC classifier three dimensions, derivational suffix; PMJ ‘three-dimensional, concave object’
7	*ha^{n}d- ~ *hat-		*j-**aɾ** PJ (RV 559)	PC ‘hand’; PJ ‘wing’
8	(*tak)		*tak PNJ (NS)	PC ‘to hit with a roundish object’, PNJ ‘to hit’
9	*^{n}da^{ŋ}g		*^{n}dap PNJ (NS)	PC ‘salt’, PNJ ‘sour’
10	*^{n}daʔ ~ *taʔ	*tɛ̃(C) (Na 55); *jɛ̃ ~ *tɛ̃	*tɛ̃ PCJ (Nb 299)	PC ‘to go’, PMJ, PCJ ‘to go, come (sg.)’ (further grammaticalization into a marker of future tense, dative/allative)
11	*^{n}da ~ *ta	*jɛ ~ *tɛ		PC, PMJ locative
12	*^{ŋ}g^{w}aʔ	*^{n}gɾɛ(C) (Na 58)	*^{n}gɾɛ PJ (Nb 294), PJ *ŋɾɛ (RV 559)	PC ‘child’; PMJ, PJ ‘egg’
13	*taʔBa		*^{n}dɛp PNJ (NS)	PC, PNJ ‘ripe’, cf. also PNJ *təm ‘raw’ (NS)

14	*^{m}ba^{n}di ~ *^{m}bai^{n}d		*mɛ̃ PNJ (NS)		PC ‘human being’, PNJ ‘people’
15	*^{m}ba^{n}d-		*mɛ̃ PNJ (NS)		PC ‘all’, PNJ plural
16	*^{m}ba^{n}d-		*^{m}bɛɲ PNJ (NS); *^{n}bɛni PJ (cf. Nb 294)		PC ‘tasty’, PJ, PNJ ‘honey’
17	*La			Apãniekrá *ɾɛ* ‘small’ (Castro Alves 2004: 48), Xavante *-ɾe* diminutive (Lachnitt 1988: 38)	PC ‘egg, offspring’
18	*apa	*ap ~*aw ~ *?o			PC ‘body’; PC? and PMJ generic object prefix
19	*hau^{m}b ~ *a^{m}bu?		*ç**ɔ**/*j**ɔ** PJ (Nb 289)		PC, PJ ‘to bathe’
20	*^{n}da^{m}ba			*ɾɔp^{ə} (Na: 56)	PC ‘feline’, PNJ ‘jaguar, dog’
21	*tau			*ʈɔ ~ *ɖɔ (Na: 26, 56)	PC ‘dog’, PNJ ‘dog, fox’
22	*^{n}da^{m}ba		*tɔ̃ PNJ (NS)		PC, PNJ ‘brother’

23	*^{m}ba	*mõŋ (Na 55)	*mõ / *mõɾ PNJ (NS); *mõ PCJ (Nb 297)	PC ‘to go’, PMJ, PNJ, PCJ ‘to go, come (pl.)’ (further grammaticalization into a marker of future tense, purposive / dative / allative)
24	*ki	*ki		locative
25	*^{n}di	*ji ~ *ɾi		PC locative; PMJ locative/instrumental
26	*hi ~ *iʔ	*ij		PC proximate demonstrative; PMJ, PJ 1st person
27	*^{n}dii(k)	***ʃĩ̃j**a(C) / ***ɲĩ̃j**a(C) (Na 71)	***j-ĩ̃j**a PJ (RV 569); ***ɲĩ̃ja**-kɾɛ PNJ (NS)	PC, PMJ, PJ, PNJ ‘nose’
28	*^{ŋ}gu		*bɯ PJ (Nb 293), *bɯ / *bɯɾ PNJ (NS)	PC, PJ, PNJ ‘take’
29	*u^{m}ba	***ʃɯm** (Na 74)	***çɯb^{n}** PJ (Nb 293)	PC ‘face, eye, fruit’, PMJ, PJ ‘seed’
30	*kuHndi ~ *kuiHnd ~ *BuHndi		*^{m}but PNJ (NS)	PC ‘throat, hole’, PNJ ‘neck’

31	**^{n}du(^{n}d)**		*tut (NS)	PC ‘bird, dove’, PNJ ‘pigeon’
32	*^{n}du		***j-u**m PJ (RV 558)	PC ‘father, uncle, ancestor’, PJ ‘father’
33	*^{n}duʔ	***ʃɔ**j / ***jɔ**j (Na 79)	***j-u**a PJ (RV 569)	PC, PMJ, PJ ‘tooth’
34	*^{n}d(i)-	*j-, *ɲ- (Na 7)	*j- PJ (RV 551)	PC relational element$_1$, PJ linking prefix
35	*ʦ(i)-	*ʃ- (Na 7)	*s- PJ (RV 551)	PC relational element$_2$, PJ third-person marker

My own reconstructions, if not indicated otherwise.

As can be gleaned from table 125, Chibchan–(Macro-)Jê sound correspondences seem to be quite recurrent. As to consonants, Proto-Chibchan velar stops correspond with Macro-Jê bilabial stops in sets 4, 5, 9, and 28. This correspondence is not unexpected: a variation between velar and bilabial stops has also been observed in Chibchan languages (see subsection 3.3.1 above), and also in some Macro-Jê languages: compare Apinajé /ko/, /po/ 'stick' (Salanova 2001: 29) or Proto-Jabutí *ku, *pu 'to eat' (Ribeiro & van der Voort 2010: 569).

The semantic shifts implied in sets 4 and 5 need further investigation. It seems that Proto-Chibchan *ka 'leaf' (possibly also 'upper part of a tree or other plant', see subsection 2.2) is cognate with Proto-Macro-Jê *pa(C) 'arm, branch', whereas Proto-Chibchan *ka^{n}d- ~ *kat- 'stick, bone, tree' corresponds with Proto-Macro-Jê *par^{ɔ} 'foot' (there seems to be no Proto-Macro-Jê form for 'leg').[114]

A semantic equation 'dog'/'jaguar' is implied in the correspondence of 'dog' and 'feline' terms in sets 20 and 21. This is perfectly expected, given that Dixon and Aikhenvald (1999: 7–9) and Adelaar (2013: 124) argue that both terms are frequently colexified in languages of Lowland South America. The Chibchan and Macro-Jê forms in question should therefore be comparable. Although in shared 'dog' terms, an origin in borrowing can certainly not be excluded (cf. Pache et al. 2016), it is noticeable that Chibchan and Northern Jê languages seem to share *two* terms for dogs or related animals, not just one: Proto-Northern Jê *ɾɔp^{ɔ} 'jaguar, dog', *ʈɔ ~ *ɖɔ 'dog, fox' versus Proto-Chibchan *^{n}da^{m}ba 'feline$_1$', *tau 'dog'. (Whether or not the Proto-Chibchan terms *^{n}da^{m}ba 'feline$_1$' and *tau 'dog' are ultimately etymologically related still needs to be worked out; the same is true for the two Proto-Northern Jê counterparts. If this is the case, there would of course be only one potential cognation set, not two.)

Table 126 shows the vowel correspondences that are attested in the presumably cognate elements of table 125.

[114] Compare also the similarity of the Proto-Macro-Jê forms in question with Proto-Cariban *a**pô**-rɨ 'arm' and ?*__pôre__(-pɨ/pa) 'leg' (Gildea & Payne 2007: 45, 58).

TABLE 126
PROTO-CHIBCHAN–PROTO-(MACRO-)JÊ VOWEL CORRESPONDENCES ATTESTED IN TABLE 125

Proto-Chibchan	Proto-(Macro-)Jê, Proto-Cerrado Jê, Proto-Northern Jê	Set
*a	*a	1, 2, 3, 4, 5, (6), 7, 8, 9
*a	*ɛ, *ɛ̃	10, 11, 12, 13, 14, 15, 16, 17
*aB, *au	*ɔ, *ɔ̃, (*o)	(18), 19, 21, 22
*a / _^{m}b; a / ^{m}b_	*ɔ, *ɔ̃	20, 23
*i	*i	24, 25, 26, 27
*u	*ɯ	28, 29
*u	*u	30, 31, 32, 33

Table 127 shows the consonant correspondences that are attested in the presumably cognate elements of table 125.

TABLE 127
PROTO-CHIBCHAN–PROTO-(MACRO-)JÊ CONSONANT CORRESPONDENCES ATTESTED IN TABLE 125

Proto-Chibchan	Proto-(Macro-)Jê, Proto-Cerrado Jê, Proto-Northern Jê	Set
*^{m}b	*m, *^{n}b, *^{m}b, *b^{n}	14, 15, 16, 23, 29
*t	*t	(8), 10, 11, ?13, 21
*#^{n}d	*j, *ɲ	9, 10, 11, 27, 32, 33, 34
*^{n}d	*#t, *t#	22, 30, 31
*^{n}d / V__V	*ɾ, *n	5, 7, 16
*k	*k, *p, *^{m}b	3, 4, 5, (8), 24, 30
*^{ŋ}g	*#b, *p#	9, 28
*k^{w}	*k^{w}, ?*kɾ	6
*^{ŋ}g^{w}	*ŋɾ	12
*ʔ	*j, Ø	10, 12, 13, 26, 33
*ʦ	*s, *ʃ	35
*h	Ø	?7, 19

Once a larger inventory of Proto-Macro-Jê forms is reconstructed, the number of cognate sets may turn out to be larger than presented in table 125, and the sound correspondences, including those of consonants, will be better understood.

In the domain of lexical parallels, a number of matches with Proto-Chibchan forms is also found in Guató, a nearly extinct language of southern Brazil (spoken in the Brazilian state of Mato Grosso do Sul, near the Bolivian border, on the banks of the Paraguai and upper São Lourenço rivers, according to Simons & Fennig 2017). Guató was classified as Macro-Jê by Greenberg (1987), Rodrigues (1999), and, with some reserve, Martins (2011). Similarities of Guató with Jê languages have also been observed by Brinton (1891a: 318), and a Macro-Jê status of this language should probably not be too hastily dismissed.[115] However, several authors (e.g., Ribeiro & van der Voort 2010; Nikulin 2015a; Jolkesky 2016) do not consider Guató to be demonstrably related to Macro-Jê, and in fact, Guató is rather different from Macro-Jê languages in terms of syntactic features and morphology (cf. Rodrigues 1999). Some Proto-Chibchan–Guató parallels are shown in table 128.

[115] The following Guató terms have remarkably similar counterparts in (Macro-)Jê: Guató *pɔ́* 'to kindle, burn' (Pal 142); *pɔ́* 'arm' (Pal 140; Pos 140); *bɔ̀* 'foot' (Pos 144), *abɔ* 'foot' (Pal 131); *nama* 'healer' (Pal 139); *ɾá* 'hand, finger' (Pal 142), 'hand' (Pos 143); *ɾékai* 'four' (Pal 147); *otɨ* 'tongue, language, speak' (Pal 141), 'language' (Pos 143) (cf. *ʧádʒà* 'tongue', ibid.); *vɛ* 'rain, to rain' (Pal 145), *vɛ̀* 'rain' (Pos 141); *pána* 'tail' (Pal 142); *kʷá* 'tooth' (Pal 138; Pos 141); *kɨ́* 'feather, hair (of head)' (Pal 138; Pos 140); *kʷɨ* 'head' (Pos 09). Most of these terms are very stable over time (cf. Holman et al. 2008) and may therefore be indicative of a potential Macro-Jê status of a part of Guató lexicon.

TABLE 128
PROTO-CHIBCHAN–GUATÓ LOOKALIKES

Set	Guató	Proto-Chibchan
1	*pána* ‘tail’ (Pal 142)	*^{m}ba^{n}d(a)- ‘tail$_1$’
2	*pèrà* (Pos 142), *pera* (Pal 142) ‘throat’	*^{m}bi-^{n}daʔ ‘throat’
3	*bɔ́* ‘smoke’ (Pal 133; Pos 142)	*^{m}ba- ‘cloud, fog’
4	*ópí* ‘red’ (Pal 140; Pos 146)	*hapi ~ *apiʔ ‘blood’
5	*ɾá* ‘hand, finger’ (Pal 142), ‘hand’ (Pos 143)	*ha^{n}d- ~ *hat- ‘hand’
6	*fṹ*, *fɨ* ‘to swim’ (Pal 134)	*hau^{m}b ~ *a^{m}buʔ ‘to bathe, swim’
7	*ópa* ‘to lie down’ (Pal 140)	*kap- ‘to sleep’
8	*ódá* ‘basket’ (Pal 139)	*kadaʔ ‘net’
9	*adá* (Pal 131), *àdá* ‘tree’ (Pos 139)	*ka^{n}d- ~ *kat- ‘stick, bone, tree’
10	*kɨ̃* ‘father’ (Pal 133)	*ka(ka) ‘father’
11	*kṹ* ‘to hear’ (Pal 138)	*kuh ‘to hear$_2$’
12	*gu* ‘to have, take’ (Pal 135)	*^{ŋ}gu ‘to take’
13	*gũ* ‘to kill’ (Pal 135)	*^{ŋ}gua ‘to kill’
14	*nama* ‘shaman’ (Pal 139)	*^{n}da^{m}ba ‘feline$_1$’

Sources are Palácio (1984), Postigo (2009) (Guató), Proto-Chibchan reconstructions are mine.

It goes without saying that some elements in table 128 are not exclusively shared by Chibchan and Guató, but also with Proto-Jê, such as the terms for ‘hand’ (set 5), ‘to bathe’ (set 6), ‘to take’ (set 12), and ‘feline’ / ‘shaman’ (set 14).

In the domain of phonotactics, there seem to be two parallels between Chibchan and Guató: First, a variation of the type *$C_1V_1C_2V_2$ and *$C_1V_1V_2C_2$ is frequently found in Proto-Chibchan (compare subsection 2.4.4.4 above), for instance in *^{m}ba^{n}di ~ *^{m}bai^{n}d ‘human being’. Most similar cases of metathesis have been observed in Guató by Schmidt (1905: 259), for instance in <tari> ‘thunder’ which, according to his analysis, would recur in metathesized form in <mu kia **tair**> ‘cloud’. Also, for Proto-Chibchan, deletion or prothesis of *(h)a in certain environments has been postulated above (compare subsections 2.4.4.2 and 2.4.4.3). In this context, it is interesting to observe that Proto-Chibchan *ha corresponds with Guató *o* in set 4 and to zero in sets 5

and 6 in table 128. However, the overall structure and the morphology of Guató and Proto-Chibchan are very different.

3.3.6 Typological parallels

Beyond the level of formal matches, several typological features are shared by Chibchan and Macro-Jê languages, above all between Chibchan and Jê proper. Typological coincidences are expected to occur in genealogically related languages (Fox 1995: 248; cf. also Dunn et al. 2005), and the present subsection will briefly mention some typological features shared by Macro-Jê and Chibchan languages, dealing with similarities in phonology and phonotactics (3.3.6.1), the morphological make-up (3.3.6.2), syntax (3.3.6.3), and semantics (3.3.6.4). Needless to say, some of the typological features mentioned below, for instance postpositions and verb-final constituent order, are typological concomitants (cf. universal 4 of Greenberg 1963). The same is true for postpositions and the genitive preceding the governing noun (cf. universal 2 of Greenberg 1963). Nonetheless, the typological matches described in the following subsections are in keeping with the genealogical relationship between Chibchan and Macro-Jê languages proposed in this study.

3.3.6.1 Phonology and phonotactics

Vowel nasality determines consonant nasality in Proto-Chibchan and in the Central American Chibchan languages Cabécar, Bribri, and probably also in Bocotá, as mentioned above (see subsection 2.4). A similar phenomenon is found in several Macro-Jê and Jê languages (Rodrigues 1999: 171–2), for instance in Kaingáng, where this phenomenon is even attested in loanwords from Portuguese: compare Kaingáng /ˈbɤ̃ko/ [ˈmɤ̃ko] ‘bench’ < Portuguese [ˈbɐ̃ŋkʊ] ‘bank’ (Jolkesky 2010: 64; more information on allophones of the voiced bilabial consonant phoneme in this language can be found in Wetzels 2010). This feature is attested in several languages of Lowland South America (e.g., of the Tupían and Tucanoan families).

Whereas no phonemic distinction seems to have existed in Proto-Chibchan and Proto-Macro-Jê between voiced oral and nasal stops, there was a phonemic distinction between voiced and voiceless stops in both proto-languages (see subsection 2.4.2 above, and Nikulin 2015a: 39–40).

Also, it has been observed that in Proto-Chibchan, the glottal stop *ʔ only occurs in syllable coda position, whereas the glottal fricative *h occurs more frequently in syllable onset. In certain cases, variation of initial *h and final *ʔ is reconstructed, for instance in *hapi ~ *apiʔ ‘blood’. This situation is reminiscent of Apãniekrá, a Northern

Jê language, where Castro Alves (2004: 37) observes that /h/ is realized as [h] in syllable onset, and as [ʔ] in syllable coda.

Finally, in terms of phonotactics, it was also pointed out above (subsection 2.4.4) that on the phonetic level, Proto-Chibchan may have had syllable-initial consonant clusters, such as those of the *CL*-type (with C as a bilabial or velar stop); these clusters occur in several Chibchan languages, most of all in Central America. They are also widely attested in (Macro-)Jê languages, for instance, in Kaingáng (e.g., Wetzels 1995: 269). Indeed, Nikulin (2015a: 99) observes that *Cɾ clusters in syllable-initial position are "very typical of Macro-Jê but disallowed in P[roto-]T[upían] and P[roto-]K[ariban]." Outside the Chibchan, Misumalpan, Macro-Jê, Jirajaran, and Yanomaman families, *muta-cum-liquida* consonant clusters in syllable onset are not frequent among the indigenous languages of Mesoamerica and South America (Willem Adelaar, p.c.).

3.3.6.2 Morphology

The morphological structure of Proto-Chibchan is relatively straightforward:

- Verbal person was probably indicated by free forms or clitics originally, not by affixes (Constenla Umaña 2012: 408); nonetheless there is a tendency in some modern Chibchan languages, to indicate object person by a prefix or clitic preceding the verbal root in transitive constructions (cf. Pache 2015: 83).
- In Proto-Chibchan, preverbal or prenominal grammatical elements were relatively rare and indicated possession and valency (cf. subsections 2.2 and 2.5.2 above).
- Proto-Chibchan had postpositions indicating case (such as instrumental, comitative or allative) (cf. subsections 2.2 and 2.5.2 above).
- It is impossible to reconstruct a Proto-Chibchan element marking the direct object in a transitive construction. Instead, different Proto-Chibchan locative/instrumental postpositions seem to have developed into agent-disambiguators or ergative markers in several Chibchan languages (see above, subsection 2.2).
- In several Chibchan languages, auxiliaries indicate aspect (e.g., Quesada 2007: 69–70, 132).

Similar observations hold for Proto-Macro-Jê. The rather simple morphosyntactic structure of this language was retained in several Jê languages and in Jabutí languages (see above, section 3.3.1). Other features, shared with Proto-Chibchan or modern Chibchan languages are the following:

- Object person, rather than subject person is indicated on the verb by prefixes in Djeoromitxí transitive constructions (Ribeiro & van der Voort 2010: 534).
- Prefixes or proclitics must have been relatively rare in Proto-Macro-Jê; they are attested in possession and valency marking (cf. Rodrigues 1999; Ribeiro & van der Voort 2010).
- In Macro-Jê languages as defined here, postpositions are used, not prepositions (Rodrigues 1999: 188–9).
- Direct-object marking is not very prominent in Macro-Jê languages (cf. Rodrigues 1999; Ribeiro & van der Voort 2010). Instead, different locative/instrumental postpositions seem to have developed into agent-disambiguators or ergative markers in transitive constructions (see above, subsection 3.3.4.1)
- Aspect and tense are frequently indicated by auxiliaries in Macro-Jê languages (Ribeiro & van der Voort 2010: 552).

3.3.6.3 Syntax

With regard to word order, I propose that adjectives and numerals followed their head in Proto-Chibchan, as argued in subsection 2.5.2 above, and since these word orders are also the most widely attested in Chibchan (cf. Quesada 2007). This order is reflected, for instance, in Muisca, as shown in (59):

Muisca (Lugo 1619: fol 3r)

(59) <cħa cħo>
man good
'good man'

In Macro-Jê languages, adjectives likewise follow the head noun (cf. Rodrigues 1999: 193). This is illustrated by an example from Panará (Northern Jê) in (60).

Panará (Dourado 2001: 34)

(60) *ĩkiej* *ikĩn*
woman beautiful
'beautiful woman', 'the women is/was beautiful'

As to possessive constructions, it is proposed here that the element indicating the possessor preceded the element indicating the possessed entity in Proto-Chibchan (see subsection 2.5.2 above). An example for a possessive construction from Cabécar is (61).

Cabécar (Margery Peña 1989: xlii)
(61) *bá* *kága* *duwá*
you father brother-in-law
'your father's brother-in-law'

Also, in the Macro-Jê languages considered here, the order is possessor–possessed (cf. Rodrigues 1999: 190–1), as illustrated by the Timbira (Northern Jê) possessive construction shown in (62).

Timbira (Rodrigues 1999: 190)
(62) *kapi* *kʰɾa*
Capi child
'Capi's child'

As to constituent order in transitive constructions, one may tentatively propose that it was SOV in Proto-Chibchan, given that this is the preferred pattern reflected in Chibchan (Quesada 2007: 78, 144). In the order of elements used for verbal person marking, both SOV and OVS are the most frequently attested in Chibchan (Pache 2015: 83). As to transitive constructions in Jê, SOV (AOV) is the most frequently attested constituent order in declarative clauses (Rodrigues 1999: 187). SOV or OVS are frequently attested as interchangeable, alternative constituent orders in transitive constructions in Jabutí languages (Ribeiro and van der Voort 2010: 525). According to Greenberg (1963), OVS is extremely rarely attested in the languages of the world. There are some few cases of SOV/OVS languages mentioned by Dryer (2013), all spoken in the Americas: Wichita (Caddoan), Apalai and Makushi (Cariban languages).

3.3.6.4 Semantics

Some Chibchan–Macro-Jê parallels are also observable in the domain of semantics. There is root suppletion in several Chibchan and Macro-Jê motion verbs in the context of verbal number marking, as in Muisca <mi> 'to be in motion (singular)' versus <gu-> 'to be in motion (plural)' (Adelaar 2005), or in Bribri [ár] 'to be hanging (singular)' versus [ˈʧĩnɨ̀k] 'to be hanging (plural)' (Chevrier 2017a: 594). A similar phenomenon exists in Macro-Jê. In Apinajé (Northern Jê), for instance, there are the forms *tẽ* 'to go (non-plural)' versus *bra* 'to go (plural)' (Oliveira 2005: 129), and for Proto-Macro-Jê, Nikulin (2015a: 55) reconstructs *tɛ̃(C) 'to come, go, walk (singular)' versus *mõŋ 'to come, go, walk (plural)'.

In other cases, suppletion in Chibchan and Jê motion verbs is found in the domain of aspect marking, as in Térraba *bí* 'to go (perfective aspect)' (Constenla

Umaña 2007: 239, 252) versus *tɔ́* ‘to go (imperfective aspect)’ (Constenla Umaña 2007: 252–3), or in Proto-Southern Jê *wɯr ‘to go (singular, perfective aspect)’ versus *tĩ(-g) ‘to go (singular, imperfective aspect)’ (Jolkesky 2010: 243). Thus, in Chibchan and (Macro-)Jê languages, suppletive forms of motion verbs exist in the domains of aspect and verbal number marking. Alternatively, these suppletive forms might be interpreted as independent roots with particular semantics. This interpretation implies a certain ‘overspecification’, as it were, of aspect and number in certain Chibchan and (Macro-)Jê roots.

In contrast with that, motion verbs in these language groups may also be semantically ‘underspecified’ with respect to another category, namely direction. In Chibchan, an example from Rama is the light verb *tiŋ* in *juk tiŋ* ‘to sit down’ (*juk* ‘to sit’) (Grinevald et al. 2002–06; cf. also Craig 1989: 80–3). Rama *tiŋ* seems to have a general meaning ‘to move in time or space’. It also occurs in constructions such as *tup tiŋ* ‘to sink’ (*tup* ‘down’) and *juwa tiŋ* ‘to become old’ (*juwa* ‘old’) (Craig 1989: 80). In Muisca lexicon, low semantic specificity of some motion verbs has been observed by Uricoechea (1871: 73) and discussed by Adelaar (2005), for instance in the case of <mi> ‘to be in motion (singular)’ or <gu> ‘to be in motion (plural)’. These roots must be combined with an unbound preposition (e.g., <hui> ‘inside, interior’, ultimately derived from Proto-Chibchan *hu ‘house’) in order to specify the direction of the motion referred to. Thus, Muisca <hui mi-> means ‘to enter (one)’ and <hui gu-> means ‘to enter (several)’ (González de Pérez 1987: 258; Adelaar 2005). Likewise, Proto-Chibchan *^{n}daʔ-u ‘to enter’ seems to contain two morphemes, *^{n}daʔ ‘to go’ and an element *u which seems to be etymologically related to Proto-Chibchan *hu ‘house’. A similar semantic ‘underspecification’ of motion verbs is found in Apinajé (Northern Jê). In this language, semantically pale motion verbs are combined with additional morphemes in order to indicate the direction of the motion referred to. The morphemes in question may encode meanings such as ‘venitive’, ‘andative’, ‘cislocative’, or ‘translocative’ (Oliveira 2005: 172). This is illustrated in (63).

Apinajé (Oliveira 2005: 172)

(63) *na* *pa* *ra* ***ma*** ***mõ***
REA 1 PFV **CTFG** **go**
‘I’m going away.’

Whereas the construction *ma mõ* is translated as ‘go away’ in (63), the meaning of *mõ* alone is translated as ‘come’ by Oliveira (2005: 172). Similar examples can be found in Apãniêkrá, likewise a Northern Jê language (cf. Castro Alves 2004: 61).

In itself, the occurrence of semantic 'overspecification' in terms of participant number or aspect, and of semantic 'underspecification' in terms of direction is probably not unusual in motion verbs. For instance, lexical encoding of participant number in verbs (most frequently verbs of motion and/or position) is also widely attested in some North American languages (Booker 1982). What is relevant in the cases discussed above may rather be the pattern of combining semantic overspecification and underspecification which is shared by some Chibchan and (Macro-)Jê languages, namely the 'overspecification' of aspect and/or number, and the 'underspecification' of direction in certain motion verbs.

Finally, a particular semantic equation – 'egg'/'child' – is found in Proto-Chibchan and Jê: Proto-Chibchan *La may both refer to 'egg' and 'offspring'; this case of colexification has been proposed for Chibchan by Constenla Umaña (1981: 370). Likewise, Kaingáng *krɛ̃* 'egg' derives from Proto-Jê *kra 'offspring' (Ribeiro 2012: 266).

3.4 Summary and discussion

The topic of section 3 was the external classification of Chibchan. Its goal was to identify, if possible, a language or language group which is demonstrably related to Chibchan in genealogical terms. The results presented in subsection 3.3 suggest the existence of a genealogical connection between Chibchan and Macro-Jê languages. The present section briefly summarizes the core evidence for this proposal (3.4.1) and discusses its implications for the prehistory of Chibchan-speaking groups (3.4.2).

3.4.1 Chibchan and Macro-Jê

The evidence in favor of a genealogical connection between Chibchan and Macro-Jê is mainly based on morphological parallels. The grammatical forms in question are short, but the fact that several of them are embedded in sets excludes the possibility that they are all fortuitous. Table 129 summarizes some of the most relevant of them: three matching sets of shared allomorphs (1 and 2; 3–5; 6 and 7), a matching set of four case-marking postpositions (8–11), and a matching set of deictic elements (12 and 13).

TABLE 129
ALLOMORPHS AND GRAMMATICAL ELEMENTS SHARED BY CHIBCHAN AND MACRO-JÊ

Set	Proto-Chibchan	Proto-(Macro-)Jê	Meaning, function
1	*^{n}da ~ *ta	*jɛ̃ ~ *tɛ̃	'to go$_1$'
2	*^{m}bã	*mɔ̃(ŋ)	'to go$_2$'
3	*a-	*a; *a*- (Apinajé)	valency-reducer$_1$
4	*aʔ-	*aʔ*- (Apinajé)	valency-reducer$_2$
5	*a^{n}d-	*at*- (Apinajé)	valency-reducer$_3$
6	*ʦ(i)-	*s-	relational element$_1$
7	*^{n}d(i)-	*j-	relational element$_2$
8	*ki	*ki	locative postposition
9	*ka	*ka	allative/dative
10	*^{n}da ~ *ta	*jɛ ~ tɛ	locative postposition
11	*^{n}di	*ji ~ *ɾi	instrumental/locative postposition
12	*hi ~ *iʔ	*iC	deictic element, proximate
13	*a	*a	deictic element, distant

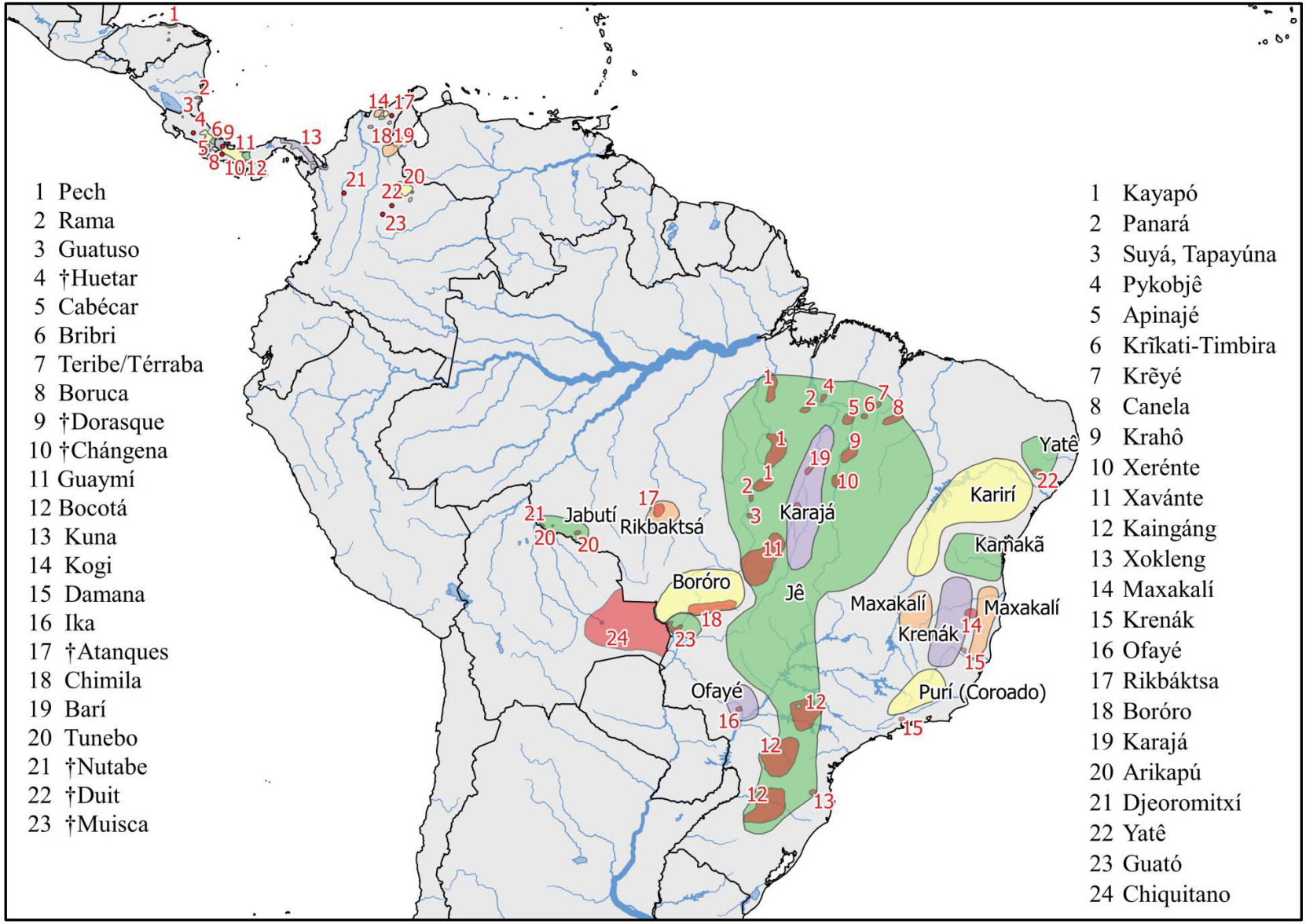

Figure 6: The distribution of Chibchan and of languages that have been classified as Macro-Jê. The map was created by Arjan Mossel, based on data from Ribeiro & van der Voort (2010: 546), Constenla Umaña (2012: 394), and Simon and Fennig (2017). Chiquitano is mainly spoken in Bolivia, but also in some adjacent areas of Brazil.

These parallels seem to indicate the existence of a genealogical relationship between Chibchan and Macro-Jê languages for the following reasons:

- There is not only a match in form and function, but also in position of the morphological elements that were treated above, for instance, in the case of the valency-reducing prefix *a- and related forms, and in the case of the generic object marker *apa and related forms, which are among the relatively few preverbal grammatical elements (that is, proclitics or prefixes) in the respective language groups.
- The pool of compared elements is a further, central argument for the validity of the proposal presented here. As mentioned in subsection 3.1.1.1.2, the higher the number of forms that are compared between two languages, the higher is the probability to find chance correspondences. However, the pools of Proto-Chibchan and (Macro-)Jê morphology are quite limited and constitute a closed class of elements.
- It is a further argument for the diagnostic relevance of the correspondences identified here that we find matches throughout the entire inventories of Proto-Chibchan and Macro-Jê morphology and not only in a single subset such as, for example, case markers.
- Some additional, lexical parallels shown in table 125 illustrate the existence of recurrent sound correspondences in the context of basic vocabulary items. Notwithstanding, more lexical cognates need to be identified for a full appreciation of Chibchan–Macro-Jê sound correspondences and, eventually, for a reconstruction of the shared ancestor language. This may be possible after the publication of Nikulin (in prep.).
- There are several parallels in terms of the overall typological profiles of Proto-Chibchan and Proto-(Macro-)Je: There is a phonemic distinction between voiced and voiceless stops, but not between voiced oral and nasal stops. Also, both proto-languages have a relatively straightforward morphology, with prefixes manipulating verbal valency, and there are several similarities in word order. Although some of these features are also attested in other Lowland South American languages (e.g., Proto-Tupían, Proto-Cariban, cf. Gildea 2012; Rodrigues & Cabral 2012), and although some of them are typological concomitants, these additional matches are perfectly in keeping with a genealogical link between Chibchan and Macro-Jê.
- The proposal of a genealogical connection between Chibchan and Macro-Jê is supported by the observation that Proto-Chibchan shares most parallels with Northern Jê languages, which are among the most conservative members of the Macro-Jê stock (cf. Adelaar 2008: 15).

- Bribri and Cabécar, two Chibchan languages of Central America, are more similar to (Macro-)Jê languages than Colombian Chibchan languages in certain respects, for example in terms of phonology (vowel nasality determining consonant nasality) and possibly also in terms of phonotactics (frequent and similar consonant clusters of the CL-type in syllable-initial position).
- Recent language contact between Chibchan and Macro-Jê languages as a source of the parallels discussed here can also be ruled out because all languages that have been proposed as Macro-Jê are presently spoken south of the Amazon.

If Chibchan is genealogically related to Macro-Jê, and if Macro-Jê is genealogically related to Tupían and Cariban languages, as suggested by Rodrigues (1985, 2000, 2009), Chibchan should also be genealogically related to Cariban and Tupían languages. In this sense, the Chibchan–Macro-Jê connection argued here would also corroborate the first external relation that had ever been proposed for Central American Chibchan languages: Herzog's (1886) proposal of a genealogical connection between Bribri and other Chibchan languages of Central America on the one hand, and Tupían (Tupí–Guaraní) and Cariban in South America on the other hand. The exact relations between Chibchan, Macro-Jê, Tupían, and Cariban remain as yet to be determined. The same would be true for the exact relations with Guaicuruan, given the fact that a genealogical connection of Macro-Jê with Guaicuruan has been proposed by Viegas Barros (2005b).

The Chibchan–Macro-Jê proposal argued here does not falsify, as it were, previous proposals of genealogical connections of Chibchan with Misumalpan, Lencan, Páez or Chocoan. For the moment, however, there is more evidence for a relatively close connection between Chibchan and Macro-Jê languages. Also, in the cases of Misumalpan, Lencan, and Chocoan, it was shown in subsection 3.2 that the arguments proposed for a genealogical link of these language groups with Chibchan are not necessarily more compelling than evidence for alternative connections of Misumalpan, Lencan, and Chocoan with non-Chibchan languages (Páez–Andakí, Taruma, Pumé, respectively). This is all the more the case since recent borrowing can by and large be ruled out as a source of matches between the latter languages (e.g., Chocoan and Pumé), but not in the case of coincidences of these languages of the Intermediate Area (e.g., Chocoan) with Chibchan. Therefore, it seems that for the moment, Macro-Jê languages may arguably be considered the closest relatives of Chibchan languages, not Misumalpan, Lencan, or Chocoan. The connection between Chibchan and Xinkan may need particular attention since Xinkan seems to have several parallels in Proto-Southern Jê basic lexicon.

Finally, the connection of Chibchan with Macro-Jê languages is perfectly in keeping with the observation of cultural and genetic similarities between Chibchan-

speaking groups or populations of the Intermediate Area and eastern South America, for instance in the domain of split greenstone pendants (see above, subsections 1.1.2 and 1.1.3).

3.4.2 Implications for Chibchan prehistory

The issue of Chibchan external classification is closely linked to questions of Chibchan prehistory. Indeed, linguistic data are a useful instrument for determining the geographical origin of a language family (Sapir 1916; Dyen 1956; Davletshin 2012; Wichmann et al. 2010), and the external classification of Chibchan has been considered an important piece of evidence for prehistoric, intercontinental relations, given that Chibchan languages are spoken on the isthmus connecting North and South America, (e.g., Mason 1950: 175; Constenla Umaña 1981: 17). Family-internal diversity suggests that Chibchan proper had a center of dispersal in Central America, as argued by Constenla Umaña (1981: 352–3, 2012: 418–9) and corroborated by Kaufman (1988) and Wichmann et al. (2010) (see subsection 1.2.2 above). Additionally, the linguistic evidence discussed in subsection 3.3 shows that the original homeland of the language family must be localized in eastern South America. There are several reasons for this conclusion:

- If Chibchan is genealogically related to Macro-Jê, as argued here, and if Macro-Jê languages are genealogically related to Tupían and Cariban languages, as argued by Rodrigues (1985, 2000, 2009), this implies a genealogical unit embracing Chibchan, Macro-Jê, Cariban and Tupían. If true, three members (Tupían, Cariban, Macro-Jê) of this genealogical unit would be spoken in eastern South America, and only one (Chibchan) in Central America. A similar argumentation holds for Chibchan in the context of the Macro-Jê–Guaicuru connection proposed by Viegas Barros 2005b. The fact that within these possible genealogical units, diversity is higher in Lowland South America, suggests an ultimate origin of Chibchan in this latter area, not in Central America or Mesoamerica. (For the reasoning underlying this ‘center of gravity approach’, compare also subsection 1.2.2 above).
- An original homeland of Pre-Proto-Chibchan in Lowland South America is further corroborated by lexical parallels of Chibchan in Tacanan identified by several scholars (see subsection 3.2.3.3 above) and in Guató (see subsection 3.3.5), and by typological and lexical parallels with other languages from this area, for instance in terms of vowel nasality (cf. Constenla Umana 1985a).

The Chibchan migration from south to north is remarkable insofar as among the indigenous peoples of the Circum-Caribbean and Central American areas, a South American origin has only been known to exist for populations of the Caribbean Islands (e.g., Adelung & Vater 1813: 377–9, 681).

Notwithstanding, it is difficult to state when, exactly, the migratory movements to the north would have taken place. A remote age of the Chibchan family – that is, first splits of the hypothetically uniform proto-language, say, some 5,000 years ago – are problematic because a connection of Chibchan with Macro-Jê languages would reach so far into the past that it would probably lie beyond the boundary of genealogical relationships demonstrable by the comparative method, if it is limited to a maximum time depth of some 4,000 to 5,000 years (cf., e.g., Beck 2017: 448).

Also, an early spread of Chibchan-speaking groups across the Intermediate Area would also leave several non-linguistic observations unexplained (see section 1): It would be difficult to bring in line with archaeological findings suggesting the spread of a particular style in gold artifacts in the Intermediate Area at around A.D. 600 in a region that overlaps with the distribution area of Chibchan languages (cf. subsection 1.1.2). Although there are no compelling reasons to connect the distribution of a particular archaeological horizon (the 'international style') with the spread of a particular language or language group (Chibchan in this case), this overlap needs to be explained. Also, the similar greenstone pendants found in the Intermediate Area and in the regions of the Tapajós/Santarém and Konduri cultures of the Lower Amazon are relatively recent, probably not much older than A.D. 300 in the Intermediate Area (cf. subsection 1.1.2). A relatively recent split of Proto-Chibchan would also be in keeping with the genetic uniformity of Chibchan speakers if this uniformity reflects a small Proto-Chibchan founding population (cf. subsection 1.1.3). Notwithstanding, given the high level of diversity which is attested in this language family, this would imply that the 'mutation rate', as it were, of Chibchan languages was (or still is) remarkably high.

While the age of Proto-Chibchan and of (Pre-)Proto-Chibchan migrations still needs to be determined, the geographic origin of Chibchan is more easily identified, as argued here, and evidence discussed in this second part of the thesis suggests that both the advocates of a Chibchan homeland in Central America and in South America were right: A Proto-Chibchan center of dispersal has been localized in Central America in previous studies, and the present work suggests that the original homeland of Chibchan must be localized in eastern South America.