Inheritance, calquing, or independent innovation?
Reconstructing morphological complexity
in Amazonian numerals

Patience Epps
pattieepps@austin.utexas.edu

Abstract
The reconstruction of morphologically complex forms offers familiar problems. As illustrated by textbook examples like Bloomfield’s seemingly reconstructable anachronism—Algonquian ‘fire-water’ for ‘whisky’—the existence of corresponding complex forms across related languages can alternatively be attributed to calquing or parallel independent innovation. This paper considers the problem of accounting for the history of complex forms in the context of the northwest Amazon, where lexical borrowing is actively resisted but calquing is rampant. Where complex forms are widely shared across related and unrelated languages, is there any hope of identifying their source, or establishing their relative age in particular groups of languages? I focus in particular on numeral terms, which reveal considerable complexity among northwest Amazonian languages. I evaluate the challenges encountered in gauging the time-depth and reconstructability of morphologically complex forms, and the criteria—comparative, typological, and geographical—that must be brought to bear in weighing more or less probable histories of complex forms.

Keywords
calquing; reconstruction; complex forms; contact; Amazon; Vaupés

1. Introduction

Morphologically complex forms present particular challenges in reconstructing the history of a language. A familiar illustration is Bloomfield’s (1946)
‘reconstruction’ of a term for ‘whisky’ in Proto-Central Algonquian, *eskoteewapoowi, literally ‘fire-water’. The expected sound correspondences are present—there is no question that the components are cognate—but the word is an obvious anachronism. How did the corresponding ‘whisky’ terms emerge among the Central Algonquian languages?

As Hockett (1948: 127) observes, the existence of a common set of morphologically complex terms across related languages—like Algonquian ‘whisky’—may be due to a variety of processes:

Each of the modern languages concerned has coined new terms for the new items of material culture, using identical (cognate) morphemes according to identical patterns of formation, in such a way that the phonemic correspondences between the whole words are perfect. But it might also be that the terms themselves date from [Proto-Algonquian] times, and that since the introduction of whisky and guns by Europeans the semantic shifts have been parallel in the various languages. Or, as European influence spread, the forms may have been invented by the speakers of one language and then borrowed by loan-translation into the other languages.1

While the challenges of teasing apart similarities due to contact, inheritance, parallel drift, and pure coincidence are encountered generally in reconstruction (see, for example, Sapir, 1921; Hock, 1991; Aikhenvald, 2007; inter alia), these problems are multiplied in the case of morphologically complex forms. The difficulty of choosing among the many possible explanations leads historical linguists to be broadly skeptical of our ability to reliably reconstruct complex forms at all. For example, Hockett (1948: 128)—in noting that our linguistic inferences about speakers’ pasts are only as good as our reconstructions—observes that “the danger is less, though probably still considerable, if the forms compared are morphologically simple than if they are compounds.” Similarly, Hock (1991: 576) stresses that reconstruction of morphologically complex forms has a low degree of confidence, and warns against “putting excessive faith in such merely possible reconstructions,” as opposed to those that can be identified as probable.

Lexical compounds, like ‘fire-water’ above, are the most common source of questions in reconstructing language histories; however, I use here the more general term ‘morphologically complex form’ in recognition of the other kinds of word-formation processes (e.g. involving class terms or other derivational

---

1 Likewise, Bloomfield (1946) observed for complex Algonquian terms like ‘whisky’ and ‘gun’ that “the meaning is modern, but the habit of formation is old.”
elements) that can present similar challenges. In fact, much of the puzzle lies in the idiomatic nature of such constructions, in that a metaphorical association is drawn between the literal meaning of the component word(s)—such as ‘fire’—and the meaning of the resulting expression. While the recurrence of such idiomatic associations across related languages is not limited to morphologically complex forms, it is far more prevalent among these than it is with simplex forms.

In light of the cautions observed by Hockett, Hock, and others, how should we evaluate the histories of morphologically complex forms, particularly among related languages? To what extent are morphologically complex forms susceptible to distinct processes of language change, in comparison to those that affect simplex forms? Are there principled ways to distinguish between calquing, parallel innovation, and inheritance? Where reconstruction is concerned, are all complex forms to be shunted equally into the ‘merely possible’ category, or are there ways to distinguish between forms for which reconstruction within a given language family is more or less probable? Similarly, where we can be certain that widespread calquing of complex forms has indeed taken place, can we identify its source—which may require demonstrating its relative age within a given group of languages?

This paper explores these questions through a series of examples from languages of the northwest Amazon, particularly the Vaupés region of eastern Colombia and northwest Brazil. These languages present a relevant test case for investigating questions of inheritance and contact: across the region, high linguistic diversity, non-contiguous distributions of language families, and relatively low levels of lexical borrowing give us tools for teasing apart inheritance and contact. At the same time, contact among some of these languages has been relatively intense, and has left its traces behind in shared grammatical structures and lexical calques.

After first providing an overview of Amazonian languages and the Vaupés region (section 2), the discussion considers several cases of morphologically complex lexical forms that are widely shared among Vaupés languages, and support the observation that lexical borrowing and calquing are profoundly different processes (section 3). The investigation focuses in particular on numeral terms, which exhibit a tendency toward morphological complexity in the region (section 4). Throughout the discussion and in the conclusions (section 5), I consider the criteria by which it may be possible to distinguish between more or less probable reconstructions of morphologically complex forms. Notably, I argue that these criteria must draw on both linguistic and extra-linguistic evidence.
2. The Amazonian context and the Vaupés region

Languages of the Amazonian basin exhibit a high degree of linguistic diversity. The region is home to some 300 languages, which belong to over fifty different genetic groupings, including many isolates (Rodrigues, 2000). The few language families that are relatively spread out across the region—most notably Arawak, Carib, Tupi, and Macro-Jê—tend to non-contiguous distributions (see Dixon and Aikhenvald, 1999; Epps, 2009a; Aikhenvald, 2012a: 32).

Despite the wide diversity among Amazonian languages and peoples, linguistic, cultural, and historical evidence suggests that interaction has been frequent and often far-reaching, leading to shared features of cultural practice, discourse, and language (e.g. Neves, 2001; Beier et al., 2002). Certain regions have experienced particularly high levels of interethnic contact and (in some cases) pervasive multilingualism, as in the Vaupés, the Xingu, and the Caquetá-Putumayo regions (e.g. Stenzel, 2005; Fausto et al., 2008; Seifart, 2011). Nevertheless, recent work suggests that interactions among native lowland languages have tended to yield relatively low levels of lexical borrowing, even in regions where relatively profound contact has taken place (e.g. Carlin, 2006; Bowern et al., 2011; Seifart, 2011; Seki, 2011; Epps, 2012).

This paper takes as its primary focus the languages of the Vaupés region, located on the border of eastern Colombia and northwest Brazil (see Figure 1). This region is home to languages belonging to four distinct families: Arawak (represented primarily by Tariana, with Baniwa and several other languages on the margins); East Tukanoan (a shallow grouping of over a dozen languages, including Tukano, Desano, Kotiria/Wanano, etc.); Nadahup (represented by Hup and Yuhup, with Dâw on the margins and Nadëb further to the east); and Kakua-Nukak (with Kakua inside the Vaupés and Nukak to the west). In addition, linguistic newcomers to the region are Nheengatú (also known as Lingua Geral, a Tupi-Guaraní language spread by Jesuit missionaries in the 16th-19th centuries), Portuguese, and Spanish.

Vaupés multilingualism is striking not only on a regional level, but on an individual level as well. For the majority of the region’s ethnic and linguistic groups, adults are able to speak and/or understand multiple languages. In the case of the East Tukanoan and Tariana peoples—who are river-dwellers with a focus on farming and fishing—this multilingualism is fostered by the practice of linguistic exogamy, by which marriages are explicitly required to occur

---

2 Nadahup and Kakua-Nukak have previously been assumed to form a single ‘Makú’ family; see discussion below.
across linguistic/ethnic lines. Thus River Indian communities are necessarily multilingual, since all the married women are representatives of other language groups. The Nadahup and Kakua peoples, on the other hand—who traditionally live in the interfluvial zones and rely heavily on hunting and gathering—do not practice linguistic exogamy. However, their long-term and semi-institutionalized socioeconomic interactions with the River Peoples have led to pervasive bilingualism among adults, mostly in Tukanoan languages—although this bilingualism is not reciprocated due to the social imbalance that pertains between the River and Forest Peoples of the region (for further discussion, see Epps and Stenzel, 2013).

Multilingualism in the Vaupés comes along with a strong cultural constraint to keep the languages distinct, and to avoid mixing. For the River Peoples, this constraint is explicitly associated with linguistic exogamy; for example, Jackson (1983: 170) records a Barasana speaker’s rhetorical question, “If we were all Tukano speakers, where would we get our women?” Accidental slips resulting in code-switches or nonce borrowings are subject to social disapproval and active ridicule (see also Aikhenvald, 2002a, 2003; Stenzel 2005). Similar constraints against language mixing are in place among the Forest Peoples;
although these groups do not participate in the exogamy system, their attitudes likely derive in part from their participation in the wider Vaupés regional ‘system’. Linguistic exogamy thus offers an obvious cultural explanation for the avoidance of language-mixing in the Vaupés; however, this exogamy may itself represent a relatively extreme instantiation of more widespread regional trends, which favor an ideal of linguistic purity and stress an association between linguistic and social identity (see references above).

Vaupés multilingualism, coupled with conscious efforts to avoid mixing languages, has had particular linguistic outcomes. In general, intrasentential code-switching is not tolerated, and lexical borrowings are very few across the region’s languages—especially in light of the intensity of contact. Contrary to cross-linguistically typical patterns of borrowing, lexical loans that do exist exhibit a high percentage of bound verb roots that appear in complex compound constructions. This tendency is probably motivated by the fact that bound forms are more easily smuggled in without notice, and thus escape censorship (see Epps, 2007: 285; Aikhenvald, 2012b).

In contrast to the limited effects of contact on Vaupés lexicons, multilingualism has fostered extensive diffusion of grammatical structures and categories. As with the borrowed verb roots, this diffusion has doubtless been facilitated by a lower degree of awareness of such structures on the part of speakers. The effects of contact on Vaupés grammars are varied and pervasive, and include the development and/or restructuring of nominal classification systems, with recurrent patterns prioritizing shape and gender in classifying inanimates and animates; development of common strategies in verb serialization; the elaboration of multiple, closely comparable categories of evidentiality; etc. (see, for example, Gomez-Imbert, 1996; Aikhenvald, 2002a; Epps, 2007, 2008a; Bolaños, 2010; Stenzel, 2013). In general, East Tukanoan languages appear to have been the primary driver of contact-related change in other languages of the region, most notably Tariana (see Aikhenvald, 2002a, etc.) and Hup (see Epps, 2007, 2008a). While Tukanoan languages appear to have themselves experienced significantly fewer effects from their neighbors (although exceptions exist, such as the effects of Baniwa on Cubeo nominal classification, as described by Gomez-Imbert, 1996), the extent of family-internal borrowing is difficult to evaluate because the languages are so closely related (see Stenzel and Gomez-Imbert, 2009; Chacon, forthcoming).

The restructuring of grammatical systems among the Vaupés languages has given rise to a relatively consistent typological profile across the region, involving prosodic tone and nasalization, nominative-accusative alignment, verb-final constituent order, a strong preference for suffixing, and similar patterns of case marking and tense-aspect-mode specification, as well as the
categories listed above. For some languages, externally motivated change has led to profound grammatical differences between them and their sister languages. This is particularly evident in the Nadahup family; Hup, Yuhup, and to some extent Dâw (within and on the periphery of the Vaupés) are in striking contrast to their sister Nadëb, which is toneless, ergative-absolutive, predominantly prefixing, and exhibits only minimal evidential distinctions and nominal classification.

On the other hand, the fact that vocabulary in the Vaupés is relatively resistant to replacement by loanwords makes genetic relationship reasonably easy to establish. Nevertheless, prior assumptions of a genetic relationship between Kakua-Nukak and Nadahup, which were lumped together to form the almost certainly spurious ‘Makú’ family (see, e.g. Campbell, 1997; Martins and Martins, 1999), illustrate the danger of making uncritical assessments of genetic relationship in situations of profound language contact. Proposals for a relationship referred to a small number of similar lexical forms shared between Hup and Kakua—most of which are almost certainly loanwords; proposals were undoubtedly also influenced by the many grammatical and phonological similarities among these languages (most of which are common to the Vaupés languages generally), and by the shared cultural category of ‘Forest Indian’ (Bolaños and Epps, 2009).

3. Morphologically complex forms in Vaupés languages

3.1. Calquing in the Vaupés

A notable feature of the common Vaupés linguistic profile is the many close parallels among morphologically complex forms in the region’s languages. These shared lexical structures point to extensive calquing, which appears to have proceeded unchecked despite the general ban on the borrowing of lexical and morphological forms. As calques, these forms are distinguished by the close translation of their component parts from language to language, such that the actual morphological material remains native to the language in question, but a shared meaning arises via language contact. As noted above, a distinguishing feature of these calques tends to be their idiomatic quality, involving a metaphorically motivated semantic shift. A comparable process can thus affect morphologically simplex as well as complex forms, but complex forms are far more prone to undergo it. Examples of idiomatically motivated semantic shift in simplex forms in Vaupés languages include the use of a single term for ‘deer’ and ‘manioc tripod’ and for ‘sun’ and ‘moon’
(e.g. in Hup, Epps, 2007: 285; and Tariana, Aikhenvald, 1996: 98), and can also be seen in certain ethnonyms and place names (see 3 below).

Examples of parallel complex terms in Vaupés languages are prevalent in several semantic domains. One of these is ritual and material culture, as illustrated in (1). Many of these terms pertain to non-native material objects and other entities (e.g. ‘wristwatch’ and ‘hospital’) and owe their complexity in part to their recent innovation; others (e.g. ‘Bone-Son’, a culture figure) are probably older. Note that the closeness of the calquing varies; for example, both Tukano and Tariana include a ‘round object’ classifier in their term for ‘clock, watch’, but the Tariana variant also involves a relativizing prefix. Similarly, the culture figure is termed ‘Bone-Son’ in both Hup and Tukano, but the Tariana variant (and the Baniwa as well) is ‘One on the Bone’ (see Hill 2009 for discussion of this figure and its name). Thus it is their idiomatic nature, rather than their structure per se, that gives the clearest indication that language contact has played a role in their formation.

(1) Ritual and material culture terms (Ramirez, 1997; Aikhenvald, 1996: 98, 2002a: 229, 2003: 13; Epps, fieldnotes)

<table>
<thead>
<tr>
<th>Tukano</th>
<th>Tariana</th>
<th>Hup</th>
</tr>
</thead>
<tbody>
<tr>
<td>imiko keo-ga</td>
<td>ehkwapi ka-wa-ka-da</td>
<td>wag tæʔkéy</td>
</tr>
<tr>
<td>(day measure-cl:round)</td>
<td>(day rel-mark-th-cl:round)</td>
<td>(day measure.nmlz)</td>
</tr>
<tr>
<td>ɨhko-wi’i</td>
<td>di-tape-dapana</td>
<td>yɔb mɔy</td>
</tr>
<tr>
<td>(medicine-house)</td>
<td>(3sg.nf-medicine-cl:house)</td>
<td>(medicine house)</td>
</tr>
<tr>
<td>őä-ki</td>
<td>yapi-riku-ri</td>
<td>ɡ’æg tæb</td>
</tr>
<tr>
<td>(bone-son)</td>
<td>(bone-loc-rel)</td>
<td>deity/culture figure</td>
</tr>
</tbody>
</table>

|                            |                       | ’day measure(r)’ = clock; watch’ |
|                            |                       | ’medicine-house’ = hospital’ |
|                            |                       | deity/culture figure = ‘Bone Son’ or ‘One on the Bone’ |

Flora-fauna terms, primarily referring to species or varietal levels, also tend to involve compounding and make idiomatic reference to other entities, as seen in (2). A comparative study of Hup and Tukano indicates that roughly 80% of such binomial flora-fauna terms are direct translations between the two languages.3

3 Unfortunately, the available materials on other languages of the area do not give enough detail about species/varietal-level distinctions to determine how widely these patterns pertain. However, Aikhenvald’s (2010) Tariana dictionary includes a few terms that parallel those found in Hup and Tukano; e.g. wesiri deri ‘inajá banana’ (inajá is a palm species, Attelea maripa).
(2) Flora-fauna terms (Epps, fieldnotes; Ramirez, 1997)

<table>
<thead>
<tr>
<th>Tukano</th>
<th>Hup</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>yamā ɨ'sē</em></td>
<td><em>mõ̊ś̊y pɨt</em></td>
</tr>
<tr>
<td>(deer Amazon.tree.grape)</td>
<td>(deer Amazon.tree.grape)</td>
</tr>
<tr>
<td><em>akō dasê</em></td>
<td><em>dêh czkw'įt</em></td>
</tr>
<tr>
<td>(water toucan)</td>
<td>(water toucan)</td>
</tr>
<tr>
<td><em>mahâ mere</em></td>
<td><em>yâk min</em></td>
</tr>
<tr>
<td>(macaw ingá)</td>
<td>(macaw ingá)</td>
</tr>
<tr>
<td><em>ikti obô</em></td>
<td><em>gõ̊b pîbît</em></td>
</tr>
<tr>
<td>(inajá banana)</td>
<td>(inajá banana)</td>
</tr>
</tbody>
</table>

Place names and ethonyms also exhibit extensive parallels throughout the region, including their variants in Nheengatú (which have also become the standard forms in the regional Portuguese, as well as in the wider linguistic and ethnographic literature):

(3) Place names and ethnonyms (Aikhenvald, 1996, 2002b; Ramirez, 1997; Epps, fieldnotes; Floyd, 2013)

<table>
<thead>
<tr>
<th>Nheengatú / Portuguese</th>
<th>Tukano</th>
<th>Tariana</th>
<th>Hup</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Tukano</strong></td>
<td><strong>Dase-á</strong></td>
<td><strong>Yâse-ne</strong></td>
<td><strong>Czkw'ît-d'ôb</strong></td>
</tr>
<tr>
<td>(toucan)</td>
<td>(toucan-ANIM.PL)</td>
<td>(toucan-ANIM.PL)</td>
<td>(toucan-ANIM.PL)</td>
</tr>
<tr>
<td><strong>Pira-Tapuyo</strong></td>
<td><strong>Wa'i-kîhi</strong></td>
<td><strong>Kaphé-mene</strong></td>
<td><strong>Hôp-d'ôb</strong></td>
</tr>
<tr>
<td>(fish-people)</td>
<td>(fish-people)</td>
<td>(fish-ANIM.PL)</td>
<td>(fish-ANIM.PL)</td>
</tr>
<tr>
<td><strong>Yawarete</strong></td>
<td><strong>Yâi-poêwa</strong></td>
<td><strong>Yawhipani</strong></td>
<td><strong>Yaʔam-hûh</strong></td>
</tr>
<tr>
<td>(jaguar)</td>
<td>(jaguar-rapids)</td>
<td>(Yawi-hipani)</td>
<td>(jaguar-rapids)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Similar parallels can be seen in many other constructions in the Vaupés languages, spanning lexicon, grammar, and discourse. For example, the regional equivalent of ‘good morning’ is the following:

(4) Morning greeting:

<table>
<thead>
<tr>
<th>Tukano: <em>wā'ka-ți mi'î</em></th>
<th>Tariana: <em>kawbi-tha phia</em></th>
<th>Hup: <em>cauôät-aʔ am</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>be.awake-INTERROGATIVE</td>
<td>be.awake-visual.past.INTERROGATIVE</td>
<td>be.awake-DYNAMIC</td>
</tr>
<tr>
<td>you</td>
<td>you</td>
<td>you</td>
</tr>
</tbody>
</table>

There is no doubt that most of the structures discussed here have involved calquing among the Vaupés languages. This process tends to be easily identified where the languages belong to different families (compare Campbell et al.’s 1986 discussion of the ‘historicist’ approach to establishing that language contact has occurred). For example, Aikhenvald (2002a: 229) argues for calquing in Tariana based on the absence of comparable forms in Baniwa, Tariana’s sister language located just outside the Vaupés; similarly, some Hup examples have no attested parallels in Dâw or Nadëb.⁴ A Tukanoan source for many of these lexical constructions is consistent with the general direction of diffusion in the region (see section 2 above); on the other hand, the presence in Baniwa of a comparable culture-figure name (ñapirikuli ‘One on the Bone’) suggests that at least this term may have originated in Arawak (though it probably entered Hup via Tukano; see Epps, 2009b).

Particularly in cases where contact may not have been so pervasive as among the Vaupés languages, however, we must also consider the possibility of independent innovation—where the processes of word formation have followed parallel paths in more than one language, but without direct copying from one language to another. Of course, ‘independence’ of innovation is something of a relative notion, in that a particular change may itself be facilitated by shared structural patterns or by cultural norms and preferences that favor certain conceptual associations. Determining the likelihood of contact as a direct motivation for the shared feature essentially comes down to gauging the naturalness of the change, as is true for many other aspects of language evaluated in historical work (sounds, morphological patterns, etc.). For the complex lexical constructions considered here, the question of naturalness relates primarily to the extent to which the new meaning is idiomatic and thus more arbitrary, as opposed to being simply the sum of its parts, as well as to the particular components involved and the order of their occurrence.

Where a parallel complex form is represented among related languages, as in the case of East Tukanoan, inheritance is also a possible explanation for its presence. Yet where these languages are themselves in contact—as is true for many of the East Tukanoan languages, whose speakers regularly intermarry—it may be close to impossible to distinguish inheritance from family-internal calquing. Similarly, where the meaning and form of the complex term are relatively natural, it is difficult to distinguish contact-driven from parallel

---

⁴ For example, Hup, Yuhup, and Dâw – but not Nadëb – use a single word for both ‘sun’ and ‘moon’, as do other Vaupés languages, and Hup and Yuhup (but not Dâw or Nadëb) use ‘one hand’ for ‘five’. Unfortunately, however, lexical data from Dâw and Nadëb are scarce.
independent innovation, especially when this coincidental innovation has involved cognate components and inherited concatenation processes. Accordingly, among closely related sister languages, especially those like East Tukanoan that have remained in contact, we may only be able to establish a complex form’s history with any confidence in the case of certain calques, when these exhibit a clearly recent meaning and a relatively arbitrary collocation—such as ‘day-measure’ for ‘clock/watch’—that allow us to rule out inheritance and parallel independent innovation with some degree of certainty.

3.2. Calquing and lexical borrowing as distinct processes

In the Vaupés languages, as we have seen, morphologically complex forms are highly susceptible to calquing even where lexical borrowing is tightly constrained. This distinction indicates that, for speakers, the full transfer of form and meaning evident in loanwords involves different processes than those responsible for the transfer of meaning and some aspects of formal structure that we see in calquing.

What are the reasons behind this difference? Part of the answer undoubtedly has to do with the degree to which speakers are aware of different components of their language. As Silverstein (1981) observes, lexical forms (as linguistic elements that are relatively referential, segmentable, and presupposing) are more accessible to speaker awareness than are grammatical structures and categories. Conscious constraints against language mixing will therefore affect the lexicon far more than they affect the grammar—as we see in the Vaupés, where lexical borrowing and grammatical diffusion have proceeded at very different rates. Presumably, the fact that calquing does not involve the direct transfer of (phonological) form allows it, like grammatical diffusion, to operate largely below the level of speaker awareness.

However, the ‘limits’ of speaker awareness are clearly not represented by a discrete boundary, but rather conform to something more like a sliding scale by which particular linguistic features will be more or less susceptible to conscious manipulation, and where variation will occur both within and across individuals. In the Vaupés, verb roots are apparently less salient to speakers than are nouns (probably because, in Silverstein’s terms, they are somewhat less segmentable and less referential), and are thus more susceptible to borrowing—though still generally resistant. Similarly, speakers’ awareness of lexical calques as associated with another language is likely below that of loanwords, which involve more foreign material, but higher than that of grammatical structures, which are more function-oriented. We note, for example, that some Hup speakers identify particular calques as deriving from
Tukano, much as they do for occasional semi-established loanwords, but rarely if ever for grammatical constructions.5

The susceptibility of morphologically complex forms to contact-driven transfer is likely not only a function of speaker awareness, but also comes from the relatively active role that morphologically complex forms play in organizing systems of knowledge through discourse. In contrast to simplex lexical items, complex forms often contain extra information, such as explicit indicators of category membership, that have discursive relevance. A particularly interesting example is found in the role that flora and fauna terms—and their taxonomic organization—play in incantations across various languages of the Vaupés region and beyond (see Hill, 1988 for Baniwa-Kurripako; Buchillet, 1992 for Desano). Typically, incantations in the region invoke exhaustive catalogues of plants or animals that are seen to be associated in some way with a particular malady. A full catalog—down to the varietal level—is important to the success of the incantation. To illustrate, the following segment (5) comes from a Hup incantation to dispel the ‘spines’ or ‘splinters’ by which a malignant entity sends illness into a victim. In part (5b), the incantation lists different kinds of spiny trees that grow along the rivers, varieties of uacu (Monopteryx uacu), pupunha (Bactris gasipaes), and abiu (Lucuma sp.). Part (5c) catalogs the different kinds of plants used to make poison for darts (actually or metaphorically), including varieties of paricá, a tree from which shamans derive a hallucinogen.

(5) Hup incantation for dispelling illness-bearing ‘spines’6

(a) Yinhyoʔ, yap, wah ʔát cáp ʔáb dóʔáyátáh. Túb-cám čáb dóʔáp, cadáp b’ah ʔáb dóʔáp, n’ikán hópd’ah níhanńuy, hópd’ah níhanńuy, ʔáb dóʔiéh, hópd’ah níhanńuy…

5 For example, a Hup consultant identified the varietal flora term ‘woolly-monkey ingá’ as a Tukano calque, and noted the alternative ‘true’ Hup form as ‘curly ingá’. However, as a reviewer points out, cultural conventions relating to ‘descriptive’ terms could provide a distinct motivation for calquing; the extent to which loanwords, lexical calques, and grammatical constructions actually correspond to a hierarchy of salience awaits further exploration.

6 The binomial plant names in this text are repeated here with full glossing:

déh yáb, déh yáb b’ah, déh wáhmáw b’ah, déh č’w b’ah…

water uacu water uacu splinter water abiu splinter water pupunha splinter ‘Water-uacu, water-uacu splinter, water-abiu splinter, water-pupunha splinter…’

poho č’w b’ah, č’w hó b’ah… hówig čaʔ cěhák, yellow pupunha splinter pupunha ripe/red splinter heart grab paricá ‘yellow pupunha splinter, ripe/red pupunha splinter…’ ‘heart-grab paricá,

tshóy g’c cěhák, wé cěhák, biciw cěhák, hóp cěhák

snake bite paricá tree.sp. paricá bisiw spirit paricá fish paricá ‘snake-bite paricá, we-tree paricá, bisiw paricá, fish paricá…’
With this, I list the patoá-spines (magical vehicles of illness). First I list, the sadap splinters I list, those belonging to those fish over there (toward the river's mouth).
I list those (spines/splinters) belonging to the fish…

(b) Děh yã́h, děh yã́h b'ah, děh wahnáw b'ah, děh c'w b'ah, hōp hup-ʔíh níh, hōp hup-ʔíh níh, tiníh, tīh dōʔnīp, yūp hōp hup-ʔáy tīh dōʔnīp.

Water-uacu, water-uacu splinter, water-abiu splinter, water-pupunha splinter, the fish-man's, fish-man's, his yellow pupunha splinter, ripe/red pupunha splinter, she takes it, that fish-woman takes it (to make spines).

(c) Cæ̃hæ̃kæ̃t tīh nám bīʔnīp, ciʔdaé, yámpiyaut tīh nám bīʔnīp, kōwšt tīh nám bīʔnīp… Cæ̃hæ̃kæ̃t tīh ciwūniph būk, yūwáw ʔāh tegdeʔkylūʔb'ay… Hāwøg ćāʔ cēbēk, tāb̥y g'āc cēbēk, wē cēbēk, bicīw cēbēk, hōp cēbēk. tīh nám bīʔnīp.

They make poison from paricá, from sídè grass, from yámpia plant, they make poison from chili pepper (to put on the spine/splinter)… Their paricá-cooking pot, I put out its fire… Heart-grab paricá, snake-bite paricá, we-tree paricá, bisiw paricá, fish paricá, (with these) they poison.

Interaction among Vaupés peoples is realized through discourse; incantations, songs, stories, conversational norms, and other discourse forms are shared continuously from one group to the next. At the same time, complex lexical forms that play a role in organizing this discourse and facilitating its regional relevance—like flora-fauna binomials in incantations, parallel place names and ethnonyms, and so forth—are replicated through calquing. Thus speakers can participate fluidly within regional systems of knowledge and verbal expression, while still maintaining linguistic distinctions.

The Vaupés case offers clear evidence that the processes of contact-driven transfer available to morphologically complex forms are distinct from those involved in the direct borrowing of lexical forms, i.e. loanwords. Their transfer involves different mechanisms, and has different kinds of relevance for speakers. This distinction between calquing and lexical borrowing has implications for linguistic reconstruction: Because low levels of lexical borrowing do not necessarily correspond to low levels of lexical calquing, we cannot rely on the histories of simple forms to inform our decisions about those of complex forms.

3. Complex forms and reconstructability: Vaupés numeral terms

In regions like the Vaupés where contact is intense, we may ask whether it is ever possible to reconstruct morphologically complex forms. Even where the relative arbitrariness of collocations allows us to rule out independent
innovation, is there any hope of distinguishing inheritance from calquing when languages are related? In this section, I consider the criteria that may allow us to sort out the histories of complex forms. Morphological complexity in numeral terms provides an illustrative case study.

3.1. Terms for ‘two’ and ‘three’ in Nadahup languages

It is not unusual among Amazonian languages for numeral systems to be restricted—in many cases lacking terms above 3 or 5—and morphological complexity (including historical complexity, which may be obscured over time) of low-level numeral terms is relatively common throughout the region (Epps et al., 2012; Epps and Hansen, in prep.). In and around the Vaupés region, the Nadahup languages in particular exemplify these characteristics. The existence of morphologically complex and/or etymologically analyzable terms for even the lowest Nadahup numerals contrasts with the morphologically simplex terms for 1-3 found in Tukanoan and Arawak languages, and suggests relatively recent layers of numeral innovation. As detailed in Epps (2006), Nadëb (the most distant sister in the family) has terms only for 1-3, of which at least ‘two’ is reportedly imprecise; Dâw (which groups with Hup/Yuhup) has terms for 1-3 plus terms relating to ‘even’ vs. ‘odd’ values that can be paired with gestures for 4-10; and Hup and Yuhup (the two most closely related sisters, both located within the Vaupés) have conventionalized terms for 1-5, with a further base-5 strategy (but without fully routinized terms) for numerals up to 20.7

The Nadahup terms for ‘two’ and ‘three’ exhibit close parallels across the family, as described in detail in Epps (2006).8 As can be seen in (6), the Hup

---

7 The Nadahup family tree is represented as:

8 In addition, the terms for ‘one’ in Nadëb, Yuhup, and Hup all appear to be related to demonstratives, but not to the same ones. It is likely that a demonstrative-numeral association goes back to Proto-Nadahup, but since the opposite trajectory (numeral ‘one’ > demonstrative/article) is cross-linguistically common, the source of these ‘one’ terms is not easily established. (Sources of data for Nadahup languages are the following – Hup: Epps, 2006, 2008b: 310; Nadëb: Weir, 1984: 103-104; Dâw: V. Martins, 1994: 158; S. Martins, 2004: 391; Yuhup: Ospina, 2002: 455.)
term for ‘two’ is morphologically complex (transparently so, in one variant), and both the Hup and Dâw terms are related to the respective words for ‘eye’ in those languages, although these are not in fact cognate between Hup and Dâw:

(6) Terms for ‘two’ in Nadahup languages

<table>
<thead>
<tr>
<th>Language</th>
<th>‘Two’</th>
<th>Etymology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nadêb</td>
<td>powap ‘2, 3, a few’</td>
<td>?</td>
</tr>
<tr>
<td>Dâw</td>
<td>tîm</td>
<td>tîb ‘eye’</td>
</tr>
<tr>
<td>Hup</td>
<td>kaʔap</td>
<td>‘eye-quantity’</td>
</tr>
<tr>
<td></td>
<td>koʔap</td>
<td>‘eye-quantity’</td>
</tr>
<tr>
<td></td>
<td>kaʔap-ʔap</td>
<td>‘eye-quantity’</td>
</tr>
<tr>
<td>Yuhup</td>
<td>_bap</td>
<td>?</td>
</tr>
</tbody>
</table>

As illustrated in (7), the terms for ‘three’ all derive from the complex form ‘rubber-tree seed quantity’. The seed of the rubber tree (Hevea sp.) is distinctively three-lobed and is culturally salient in the region (being associated with an edible fruit, and also made into a favorite children’s toy).

(7) Terms for ‘three’ in Nadahup languages

<table>
<thead>
<tr>
<th>Language</th>
<th>‘Three’</th>
<th>Etymology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nadêb</td>
<td>tamawîb</td>
<td>(tâ- 3sg) + ‘rubber tree seed (quantity)’</td>
</tr>
<tr>
<td>Dâw</td>
<td>mutuwap</td>
<td>‘rubber tree seed quantity’</td>
</tr>
<tr>
<td></td>
<td>mutuwîp</td>
<td></td>
</tr>
<tr>
<td>Hup</td>
<td>mòdaʔap</td>
<td>‘rubber.tree-seed-quantity’</td>
</tr>
<tr>
<td></td>
<td>mòt-ʔap-ʔap</td>
<td>‘rubber.tree-seed-quantity’</td>
</tr>
<tr>
<td>Yuhup</td>
<td>_bodîgʷ-ʔap</td>
<td>‘rubber.tree.seed-quantity’</td>
</tr>
</tbody>
</table>

These Nadahup numeral terms are conventionalized, and their morphological complexity is not necessarily fully transparent today. However, their forms do suggest that the metaphorical association between the numeral value and the item associated with it has been maintained over time. In Hup, most notably, morphologically transparent variants of the terms for ‘two’ and ‘three’ are

---

9 The probable cognacy between Nadêb *tamawîb* and the other Nadahup ‘three’ forms was not considered in Epps (2006).
attested alongside phonologically reduced forms. In my experience with Hup speakers along the Rio Tiquié, the expanded forms are not in general use, but at least some speakers are aware of them.

The possibility that speakers may maintain an association between an abstract concept and a concrete referent over time has implications for determining linguistic history: It implies that we may be able to trace the time-depth of a particular meaning, even where the form has changed. A comparable example may be seen in terms for colors, which like numerals represent an abstraction from the concrete objects that are inherently associated with those properties. Especially for cultures where color reference carries a low functional load, color terms may come about via a gradual delinking and generalization of the abstract notion of color from specific objects—but this may be a slow process (Lyons, 1999; Kay and Maffi, 1999; Levinson, 2001).

In the Papuan language Yélî Dnye, for example, the two dialectal variants of terms for ‘red’ correspond to the two different words for a species of red parrot that exist in these dialects; as Levinson points out, the co-existence of these terms suggests "that the reference to the bird is still salient, that these are partially live rather than fully dead metaphors" (Levinson, 2001: 18). We may compare this to the existence of two distinct forms for ‘two’ in Hup and Dâw, both of which mean ‘eye’, but involve non-cognate forms.

Determining the history of the complex numeral forms ‘two’ and ‘three’ in the Nadahup family requires us to weigh the relative probability of different historical scenarios. The associated meanings (‘eye quantity’ and ‘rubber tree seed quantity’) are not widely attested in numeral terms elsewhere in Amazonia or the rest of the world, suggesting that parallel independent innovation is unlikely; this is especially true for ‘three’, which furthermore (unlike ‘two’) is attested across all four members of the family. Generalized calquing is also unlikely, because comparable forms are unattested in other local languages in the region. Family-internal calquing could only have predated the dispersal of Nadahup languages, which are currently geographically separated; while Hup and Yuhup do maintain occasional contact, and there are linguistic indications that Dâw and Nadëb may have had some contact in the past (see Epps, forthcoming), there is no evidence that Dâw or Nadëb have come into contact with Hup and Yuhup since their dispersal. Thus, while we cannot rule out the possibility of early family-internal calquing or the independent innovation of ‘two’ in Hup and Dâw, inheritance appears to be a highly probable explanation for the shared existence of ‘three’ in at least some of these languages, and a somewhat probable explanation for that of ‘two’—although in the case of ‘two’ inheritance would have been limited to meaning
and metaphorical association, rather than to form (see Epps, 2006 for further discussion). 10

In sum, our reconstructions of morphologically complex ‘two’ and ‘three’ terms for Nadahup languages are arguably of more ‘probable’ than ‘possible’ status, particularly in the case of ‘three’. To achieve ‘probable’ status, however, we must draw not only on comparative evidence (attestations across the family), but also typological information (the scarcity of comparable terms in other languages), and geographical evidence (the non-contiguous distribution of the Nadahup languages).

However, the example of numeral terms also highlights an additional difficulty in distinguishing between inheritance, calquing, and independent innovation: Perhaps especially for morphologically complex terms, associations between literal and more figurative or abstract meanings may retain their salience over time, even as form is replaced. Thus we may be able to argue for reconstruction even in the absence of cognate forms, even though this normally would be taken as evidence against inheritance. Conversely, not even the lack of cognacy in form may constitute definitive evidence for calquing or independent innovation among parallel complex forms in related languages.

3.2. Terms for 4 in the Vaupés and beyond

Morphological complexity in Vaupés numeral terms goes well beyond Nadahup ‘two’ and ‘three’. In most of the region’s languages, terms for ‘four’ are based transparently on the respective words meaning ‘companion’. Terms for ‘five’ are also complex, deriving from ‘one hand’, and where additional terms exist they involve the addition of fingers and/or feet/toes (see Epps et al., 2012). Comparative evidence indicates that ‘five’ terms in Nadahup languages and Tariana are indeed calqued from Tukanoan (Aikhenvald, 2002a: 108; Epps, 2006), but the frequency of ‘hand’ as a source for ‘five’ throughout the world’s languages makes independent innovation a likely alternative explanation for their wider distribution.

We focus here on the ‘relational’ (i.e. ‘companion’-based) terms for ‘four’, which—in contrast to ‘hand’-based terms for ‘five’—are virtually unattested outside the Amazon basin, and thus typologically unusual (see Epps et al.,

---

10 Of course, our evidence for reconstructing ‘two’ goes only as far as Proto-Dâw-Hup/Yuhup, not Proto-Nadahup. Similarly, the possibility of contact between Dâw and Nadêb suggests that ‘three’ could easily have been calqued between these two languages, and thus only inherited in the Dâw-Hup/Yuhup branch as well.
2012; Epps and Hansen, in prep.). Their cross-linguistic rarity is a strong argument against independent innovation, whereas calquing is in many instances an obvious explanation for their existence. Inheritance, on the other hand, is far more difficult to prove, but would allow us to identify a likely origin point for their subsequent diffusion.

In the Vaupés, the presence of a relational ‘four’ term (and, in a few cases, comparable variants for other values, probably created via back-formation) in Nadahup and Tariana almost certainly derives from contact with Tukanoan languages, in which such forms are widely attested (as discussed below). As noted by Aikhenvald (2002a: 108), Tariana kephunipe (ka-iphu-nipe [rel-accompany-nmlz] ‘the one who has a companion, 4’) is calqued from Tukano, as evidenced by its absence in Tariana’s sister Baniwa (which instead has ri-kwa-da-ka [3SGNF-be.enough-cl:round-decl] ‘it is enough, 4’). Similarly, Nadahup languages exhibit the relational terms listed in (8), but the absence of a comparable term (or in fact any ‘four’ term) in Nadëb, as well as the intensive contact known to have taken place between Tukanoan and Nadahup, are evidence that this too is a calque; however, it is unclear whether the calquing took place independently in each Nadahup language or in a common ancestor of Dâw and/or Hup and Yuhup (see Epps, 2006).

(8) Relational numeral terms in Nadahup

- **Hup:** (hi-)bab’-ní [(fact-)accompany OR (fact-)sibling-exist]12
  ‘sibling exists; accompanied, 4’
- Upriver Hup dialect: bab’-pã [sibling-NEG] ‘no sibling, 3’
- **Yuhup:** bàhm-~ di- wāp [companion-be-quantity OR accompany-quantity]
  ‘sibling-exist quantity; accompany quantity, 4’
- **Dâw:** më̃ʔ mām’ [one brother] ‘one brother exists; 4,6,8,10’
  më̃ʔ mām’ më̃ʔ [one brother neg] ‘one has no brother; 5,7,9’

The history of relational ‘four’ in Tukanoan languages is more difficult to establish. The term is widely distributed within the family, providing additional support for a Tukanoan source in Tariana and Nadahup. Figure 2 shows all Tukanoan languages for which I have identified a term for ‘four’ in the available documentation, according to the subgrouping established by Chacon (forthcoming). Bolded language names indicate the presence of a relational

---

11 Particularly in the case of even-odd variants like those seen in Dâw, these numeral terms may be paired with gestures and often function more like a tally system than designates of exact quantities (see Epps, 2006; Epps and Hansen, in prep.).

12 Both interpretations of the Hup form are possible.
term (see 9 and 10 below), while plain font indicates that the language’s term for ‘four’ (or ‘three’, ‘five’, etc.) is not known to be relational.

As Fig. 2 illustrates, relational numeral terms are altogether absent in West Tukanoan languages. Among the East Tukanoan languages, on the other hand, relational terms are widespread in the Western (9) and Eastern (10) branches of this subgroup. Attestations of ‘four’ attested in the Southern branch are limited to Retuarã bo’ta-raka- (etymology unclear).13

(9) East Tukanoan, Western branch:

- **Barasano**: babá-ri ‘companion(s)’
  not cognate? cf. expected bahá
- **Makuna**: hijria (contains ha ‘2’; possibly idia 3)
- **Desano**: wapiki-ri ‘accompany’
  not cognate? cf. expected baha
- **Siriano**: wa’píki- ‘accompany’
  not cognate? cf. expected baha
- **Kubeo**: yówai-kiwai ‘companion-have’
  not cognate

(10) East Tukanoan, Eastern branch (subgroups I and II):

<table>
<thead>
<tr>
<th>Language</th>
<th>Term</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Tukano</strong></td>
<td>ba’pá-ritise</td>
<td>‘having companion(s)’</td>
</tr>
<tr>
<td>I Bará</td>
<td>bapá-ri</td>
<td>‘companion(s)’</td>
</tr>
<tr>
<td>Tatuyo</td>
<td>bapa-ri</td>
<td>‘companion(s)’</td>
</tr>
</tbody>
</table>

---

13 Data on Tukanoan numerals come from the following sources: Huber and Reed, 1992; Strom, 1992; Ramirez, 1997; Miller, 1999: 46-47; Morse and Maxwell, 1999; Stenzel, 2004: 185-191 and p.c.; Gomez-Imbert, 2009; Thiago Chacon, p.c.).
As seen in (9), the majority of terms for ‘four’ in the Western branch languages are relational, but the forms themselves are not cognate (Thiago Chacon, p.c.). Although the Nadahup case discussed above does suggest that inheritance of a metaphorical association, rather than a specific construction or form, is possible, the relational terms in these languages probably owe their existence to calquing (or even, as in the Barasano case, to lexical borrowing): Barasano, Desano, and Siriano are in frequent contact (mostly via intermarriage) with languages of the Eastern branch, whereas contact involving Kubeo, Retuarã, and Makuna (of which the latter two lack relational forms) is much more limited.

For the Eastern branch, a relational ‘four’ term *bapa-ri does appear to be fully reconstructable. However, the form could also have easily been internally calqued using cognate components, since these languages experience relatively intense contact via the linguistic exogamy system. Unlike Algonquian ‘whisky’, meaning gives us no additional clues, so a reconstruction of relational ‘four’ in the Eastern East Tukanoan languages can be assessed as only possible, not probable.

In fact, a wider comparative view indicates that relational ‘four’ originated beyond East Tukanoan, regardless of whether it spread by contact or inheritance within this family. Relational terms for ‘four’ (and in some cases for ‘three’, ‘five’, or other values) are widespread throughout the Amazon basin; they occur in about 25% of 195 languages surveyed, distributed across at least thirteen different families/isolates from French Guyana to Peru to Paraguay (see Epps et al., 2012; Epps and Hansen, in prep.); compare, for example, Miraña [Boran] tsanέ’náʔbέ’yáhktiši- ‘being companions to each other, 4’, Xavánte [Macro-Jê] mro po ‘with companion/spouse, 4’, etc.

Despite their widespread occurrence across Amazonian language families, relational numeral terms are not generally reconstructable beyond very shallow time-depth. In fact, the only language families in which they occur more than sporadically are Nadahup and East Tukanoan, which as discussed above can probably be attributed (at least in part) to calquing; some Jê and Macro-Jê languages, in which their wide heterogeneity in form, value (3, 4, 5), and literal meaning (e.g. ‘spouse’, ‘friend’, etc.) also indicates that inheritance is
Note that partial or imperfect calquing is quite common among the world’s languages; compare the well-known examples of ‘skyscraper’, e.g. German Wolkenkratzer ‘cloud-scratcher’ vs. French gratte-ciel ‘scratch-sky’.

The Tupi-Guaraní family presents an intriguing exception to the sporadic attestation of relational ‘four’ terms within particular Amazonian families. Relational terms are widespread in Tupi-Guarani, making inheritance appear somewhat more plausible. Moreover, if relational ‘four’ reconstructs, its time-depth and the wide geographic distribution of Tupi-Guarani languages would point to this family as a plausible source for the diffusion of relational terms into other Amazonian languages. But what evidence do we have that relational ‘four’ is inherited in Tupi-Guarani? As shown in Fig. 3 below, relational ‘four’ terms—most of which are based on the root *iru ‘accompany’ (reconstructed by Mello, 2000: 163)—are attested in languages throughout the ‘Main’ Tupi-Guarani subgroup, according to the family tree proposed by Chousou-Polydouri et al. (in prep; see also the largely comparable subgroupings proposed by Lemle, 1971 and Rodrigues and Cabral, 2002). Mello’s reconstruction of Proto-Tupi-Guaraní includes *iruni ‘four’ (2000: 163), based on the reflexes listed in (11), but in fact these languages belong only to the Guaranían subgroup of Tupi-Guaraní (see Fig. 3), with the exception of Tupinambá (which was spoken along the Brazilian coast, adjacent to Guaranían languages).

(11) Reflexes of Tupi-Guaraní *iruni (relational ‘four’; Mello, 2000: 163)

- Tupinambá: urundí (Gonzales, 2005: 126)
- Lingua Geral: iruni
- Guaraní Mbya: iruni
- Paraguayan Guaraní: iruni
- Chiriguano: i’runi
- Old Guaraní: iruni
- Tapiete: urundí (Gonzalez, 2005: 126)

Note that partial or imperfect calquing is quite common among the world’s languages; compare the well-known examples of ‘skyscraper’, e.g. German Wolkenkratzer ‘cloud-scratcher’ vs. French gratte-ciel ‘scratch-sky’.

Anchieta (1595; see also Eckart (1890 [1753-1757]: 4) gives the form monherondye (probably mõ-irundi; see comparable examples listed by Mello 2000:163 for ‘accompany’ and ‘four’ across Tupi-Guaraní languages).
Beyond the Guaranían subgroup of the family, however, relational numeral terms are not regular in form, although most do involve some reflex of *iru ‘accompany’:

(12) Other Tupi-Guaraní relational ‘four’ terms

- Guayarú: *iruanta (Schmidt, 1936)
- Kokama: *iruaka (Vallejos, 2010: 264)
- Omagua: *iruaka (Lev Michael p.c.)
- Tembé: *uzeeiru-*iruagu (Mello, 2000: 163)
- Wayampí: *iráte (Olson, 1978: 11)
- Kayabí: *irupáwe (Dobson, 1997: 125)
- Kamaiura: *mojo'irũ (Seki, 2000)
- Parakanã: *ioirugatu (Silva, 2003: 70)
- Tapirapé: *išaíri (Prança, 2007: 149-151)

(13) Non-relational ‘four’ terms in Tupi-Guaraní

- Ava-Canoeiro: *olopakatu, oikopa'te (Mello, 2000: 163)
- Urubu-Ka’apor: *tumeme (Kakumasu, 2007: 139)
- Apiaká: *mukújatu (‘two-?‘; Mello, 2000: 163)
- Araweté: *kukũĩ (Mello, 2000: 163)
- Emerillon: *momokante (‘two-focus‘; Rose, 2003: 195)
- Sirionó: *tʃeremó-ha (Mello, 2000: 163)
- Xeta: *mɔkaj mɔkaj (‘two two‘; Vasconcelos, 2008: 13)

So what is the history of relational ‘four’ within Tupi-Guaraní? The reconstructed term *irunɨ was probably inherited in the Guaranían branch, while its occurrence in Tupinambá could plausibly be due to calquing. Elsewhere in the ‘Main’ branch of Tupi-Guaraní, inheritance is plausible, but would have necessarily involved the metaphorical association between ‘four’ and ‘companion’, rather than a fully established form—much as was suggested above for the Nadahup forms for ‘two’ based on ‘eye’. This scenario is favored by Schleicher (1998: 13) in his reconstruction of Proto-Tupi-Guaraní; in noting the variability of the forms across the family, he posits that only the strategy associating ‘four’ with *iru ‘companion’ was conventionalized when the protolanguage split up.

The only plausible alternative explanation is that relational ‘four’ entered the majority of the Tupi-Guaraní languages by calquing, whether directly

---

16 Compare also Guajájará zeiruŋatu, zeirongatu (Roberts and Symes, 1936: 245; Green, 1997: 5) and Akawere tapiar ‘(companion’; see Boudin, 1966: 90).
Figure 3. Relational ‘four’ in Tupi-Guaraní languages.
(bold + underline: reflexes of *irunɨ; bold: other ‘relational’ form; plain: ‘four’ term does not appear to be relational; classification Chousou-Polydouri et al., in prep.)

Figure 4. Tupi-Guaraní languages.
(bold + underline: reflexes of *irunɨ; bold: other ‘relational’ form)
from neighboring languages or via substratum influence. However, it is clear that the term cannot be a recent calque within the Tupi-Guaraní family; as Figure 4 illustrates, the languages are widely dispersed, and have probably been so for some 1000-2000 years (Noelli, 2008). While it is possible that relational ‘four’ entered many of the Tupi-Guaraní languages via localized processes of calquing (e.g. within the Vaupés, the Xingú, and other regions), this scenario gives us no explanation for how relational ‘four’ became so widely distributed across Amazonia in the first place, especially given its relatively shallow time-depth in the majority of the language families in which it appears. Tupi-Guaraní languages, on the other hand, present a reasonable motor of spread by which the strategy could have been implanted in these far-flung regions in the first place, after which it could have been transferred among other local languages. We can probably never know whether the strategy was first innovated within Tupi-Guaraní or was calqued into the family at some upper node of the tree; however, its recurrence among these geographically far-flung languages does suggest that it was inherited through much of the Peripheral and/or Central subgroups.

In sum, comparative, geographical, and typological evidence points to the inheritance of relational ‘four’ within at least a major branch (Guaraní or higher) of the Tupi-Guaraní family. While the general lack of a comparable form of this morphologically complex numeral term challenges our reconstruction, we can nevertheless assess it as more probable than the possible reconstructions of other complex forms in the region, such as the relational ‘four’ term in East Tukanoan languages—among which intense, ongoing contact points to family-internal calquing—and ‘hand’-based terms for ‘five’ in other families, which could be due to either calquing or to independent innovation.

4. Conclusion

As can be seen in the case of Amazonian numeral terms, complex forms are subject to a variety of historical processes that are of different relevance to simplex lexical items. The Vaupés languages offer ample illustration that complex forms may be readily calqued, even when lexical borrowing is avoided—suggesting that complex forms are susceptible to different kinds of speaker awareness and manipulation. Complex forms are likewise prone to innovation via strategies like compounding that are not applicable to simplex forms. Furthermore, when complex forms are inherited, the Amazonian evidence suggests that they may follow a trajectory that is much less accessible to
simplex forms; in particular, a *metaphorical association* may be more stable than *form*, and continued salience of the metaphor over time may even restrain full lexicalization.

In light of the distinct historical pathways open to morphologically complex and simplex lexical items, complex forms offer particular challenges to our efforts to determine linguistic history. Where complex forms are attested among related languages, their reconstruction may be very difficult to effect with any degree of certainty—since corresponding forms may arise through contact-induced change or independent innovation, as well as inheritance, and inheritance itself need not require stability of form. Despite these challenges, however, this paper has argued that not all reconstructions of complex forms are equally hopeless. We do have tools to assess the relative probability of our reconstructions, most notably in typological criteria, which provide information about the likelihood of independent innovation, and geographic-historical criteria, which inform the plausibility of contact-related transfer.

**References**


Epps, Patience and Cynthia Hansen. In prep. One, two, three, brother: numeral etymologies and language contact in Amazonia. Ms.


