

On Terena (Arawakan) *-pâho* ‘mouth’: Etymology and Implications for Internal Classification*

This paper presents an etymological investigation of the Terena noun for ‘mouth’, *-pâho*, which, according to current comparative Arawakan linguistics, lacks known cognates. I show that cognates of this Terena noun exist in Mojeño, a close relative of Terena within the Bolivia-Paraná subgroup of the Arawakan family. In Ignaciano and Trinitario, the best known modern varieties of Mojeño, the cognates of *-pâho* are semantically-shifted nouns meaning ‘door’. I propose an account of the semantic and formal relations between these forms via an etymological source **-paho* ‘mouth’ and a compound noun **paho-peti* ‘door’ (lit. ‘mouth (of the) house’). This account relies on a more general pattern of noun formation in the Arawakan family, regular phonological correspondences and is consistent with modern views on the nature of diachronic metaphorical extensions in lexical semantic shift. The evidence presented and argumentation built to support this analysis adds Terena evidence to the correspondences supporting a vowel merger in the Ignaciano variety of Mojeño. Finally, I consider some implications for internal classification, advancing the hypothesis that Terena and Mojeño form a separate branch of the Arawakan family, the Achane branch, one that does not include Baure.

Keywords: Etymology; Arawakan languages; Terena; Semantic change.

1. Introduction

The Arawakan language family is routinely celebrated as the largest language family of the New World — both by its geographic spread and by the number of languages that belong in it (Kaufman 1990: 40; Aikhenvald 1999; Campbell 2012). In addition, linguists and other prehistorians attach significant importance to this language group as a potential source of privileged insight on the pre-history of South America (see for instance the collection of papers in Hill & Santos-Granero 2002 and Hornborg & Hill 2011). Nevertheless, our knowledge of the historical linguistics of the Arawakan language family remains arguably less advanced than is the case with other large groups of lowland South America, such as Cariban and Tupian (see e.g. Michael 2009; Campbell 2012). With this situation in mind, and seeking to complement more ambitious pioneering works such as Payne (1991), recent approaches to the historical-comparative linguistics of the Arawakan family have proceeded in a bottom-up manner, reconstructing from small sets of closely related languages, with a finer understanding of the phonology and the morphology of the relevant languages yet, at the same time, keeping an eye open to broader issues, such as that of internal classification (Michael 2011; Danielsen 2011; Lawrence 2014; Jolkesky 2016; Carvalho 2015, 2016a,b,c; forthcoming).

The present paper follows in the footsteps of this program. I will be concerned with providing a sensible etymological analysis of the form *-pâho* ‘mouth’ attested in Terena, a Southern Arawakan language of Brazil, within the broader background of an approach to the his-

* I am sincerely grateful to Andrey Nikulin and to Françoise Rose for enlightening comments and observations that helped improve this paper. The usual disclaimers apply.

torical linguistics of this language and the specific low-level subgroup where it presumably belongs. First, I will discuss how the historical-comparative linguistics of the Arawakan languages has so far failed to provide a sensible account for this item. After tracing this form to its earliest attested sources, I argue that it has semantically shifted cognates in the modern varieties of Mojeño, a language closely related to Terena. Evidence from early 18th century data on Old Mojeño plays a vital role in filling the gaps that relate these forms via lexical semantic shifts. In accounting for these shifts I will not only provide an account that is consistent with recent work on the directionality of semantic change — and on how reliably identified directionality trends can help in semantic reconstruction — but I will also discuss evidence for other developments in the history of Terena phonology and morphology, thus illustrating the inherent feedback between particular etymologies and a more general understanding of language's diachrony (Mailhammer 2014: 424–425).

Data for this study comes from the author's fieldwork on Terena (Cachoeirinha Reservation, Mato Grosso do Sul, Brazil), from published sources on the modern languages (Ekdahl & Butler 1969, 1979 on Terena; Ott & Ott 1983 on Mojeño Ignaciano; Gill 1957, 1970 on Mojeño Trinitario; Danielsen 2007 on Baure) and from a 18th century grammar and vocabulary of Old Mojeño (Marbán 1702). Additional sources on other languages discussed in the paper will be referred to accordingly.

The paper is organized as follows: Section 2 offers a brief overview of the Bolivia-Paraná Arawakan languages, a subgroup of the Arawakan language family including Terena, Mojeño, Baure and other less well-described languages. Section 3 reviews the treatment of the form for 'mouth' in Terena in the existing historical-comparative literature and thus sets the issues to be tackled in the remainder of the paper. Section 4 constitutes the core of the present contribution. Finally, section 5 considers some implications for internal classification of the findings discussed here, advancing the hypothesis that Mojeño is the closest relative of Terena within the family and that both, perhaps with the inclusion of Paunaka as well, form a separate subgroup, the Achane branch.

2. The Bolivia-Paraná Arawakan Languages: Brief Overview

A branch of the Arawakan language family composed of Terena (including under this label other geographically or chronologically-defined varieties such as Kinikinau and Guaná; see Carvalho 2016a), spoken in southwestern Brazil, the Mojeño varieties (such as Ignaciano and Trinitario) and Baure, all spoken in Bolivia, is usually identified in classifications of the Arawakan languages (see Matteson 1972: 186–192, who does not include Mojeño; Kaufman 1994; Payne 1991: 489; Aikhenvald 1999: 67; Walker & Ribeiro 2010: 3; Campbell 2012: 75). Other, less well-attested languages such as Paunaka and Paikoneka are also assigned to this subgroup by other researchers (Danielsen 2011; Jolkesky 2016). The label *Bolivia-Paraná* was proposed for this group by Payne (1991) and I will retain this use in the present paper.

As is the case elsewhere in the Arawakan family, the evidence for recognizing a Bolivia-Paraná subgroup has been less than compelling, however. Areal-geographic factors have had a major influence in promoting the reality of this subgroup, as implied by its label (see specially Aikhenvald 1999). Strictly linguistic evidence has been presented by Payne (1991), but this amounts to figures for shared lexical retentions and loosely defined phonological outcomes, such as a 'weakening' of Proto-Arawakan (henceforth PA) dorsal stops (Payne 1991: 440), which, even if defined in more precise terms, are known to have occurred independently in many Arawakan subgroups. Other studies, both preceding (e.g. Matteson 1972)

and following (e.g. Walker & Ribeiro 2010) Payne's work have applied similar methods, relying on shared proportions of lexical cognates to recognize this as a coherent subgroup of the family.

As to the internal organization of the Bolivia-Paraná subgroup, there seems to be a widespread, recent opinion — again, based on geography, shared lexical retentions or other assessments of 'structural similarities' — that Baure and Mojeño are closer to each other than any of these is to Terena (Walker & Ribeiro 2010: 3; Danielsen, Dunn & Muysken 2011: 185). This claim should be taken with caution, however, not only due to the different methodologies and datasets employed in these studies, but also because no investigation has so far established this by identifying shared innovations and because other authors refrain from proposing any internal structure to this branch (see e.g. Aikhenvald 1999: 67; Danielsen 2011). The latest and most detailed study of these languages, Jolkesky (2016), does not consider Terena data, taking it as a premise that Mojeño, Paunaka, Paikoneka and Baure are more closely related to each other than any of these is to Terena. This generalized but not consensual view is expressed in the arrangement below:

- (1) Prevalent Internal Classification of the Bolivia-Paraná Subgroup:
Bolivia-Paraná Subgroup
 - Terena
 - Mamoré-Guaporé branch
 - Baure-Paikoneka
 - Baure
 - Paikoneka
 - Mojeño-Paunaka
 - Paunaka
 - Mojeño (Ignaciano, Trinitario)

I will assume the structure above as a kind of null hypothesis enshrining some common ground reflecting a certain degree of agreement that has been reached among researchers, even if strong, compelling evidence for it (especially for the classification of Terena) is yet to appear in print. My own opinion is that the existence of a Bolivia-Paraná subgroup is very plausible and perhaps even obvious; yet, unless this impressionistic and intuitive assessment is moved from the level of a gut feeling to a detailed, methodologically sound understanding of the historical development of these languages, hardly any progress can be achieved on this specific issue and on Arawakan historical linguistics more generally. Moreover, unless a detailed understanding of the diachrony of these languages is offered, there is little hope that the more puzzling question of the internal classification of the Bolivia-Paraná languages can be properly addressed. I turn to this issue in section 5.

3. On Terena *-pâho* 'mouth': Treatments so far and statement of the problem

Payne (1991: 413) presents three separate cognate sets for the meaning 'mouth', each ascribed to a different Proto-Arawakan (PA) etymon. Interestingly, however, not a single member of the Bolivia-Paraná subgroup features in these cognate sets. Since Payne's (1991) study remains to this day the most extensive historical-comparative investigation of the Arawakan languages, it is perhaps surprising that this gap has not drawn any attention in the comparative literature so far. This has additional significance as the lack of a cognate in these languages —

otherwise thought to constitute a distinct subgroup in itself — might indicate the existence of lexical or morphological innovations.

Earlier, Matteson (1972: 186–192) proposed a reconstruction of ‘Proto-Shani’, the common ancestor of Terena, Kinikinau and Baure (she did not include any Mojeño data). In the case of her reconstructed etymon for ‘mouth’, Matteson (1972: 191) postulates a form **báaho* based on Terena and Kinikinau reflexes alone. Setting aside the fact that Terena and Kinikinau are close dialectal variants of the same language, so that cognate matches involving only these two speech varieties could hardly justify reconstructing an etymon at the deeper level of her Proto-Shani, there are many shortcomings in Matteson’s data, particularly with respect to Kinikinau. The differentiation between Kinikinau and Terena has been mistakenly overestimated as an artifact of poor morphophonological analysis: the word-initial voiced stop in **báaho*, for instance, results from an incorrect analysis of Kinikinau as having phonemic voiced stops (see Carvalho 2016a for details on these and other points and Payne 1991: 368–371 for a general evaluation of Matteson’s 1972 study). Be as that may, the form **báaho* reconstructed by Matteson does not appear in her Proto-Arawakan cognate sets, thus underscoring the isolated status of the Terena noun *-pâho*.

Jolkesky (2016: 13) notes the existence of non-cognate material for the meaning slot ‘mouth’ in a comparison of Mojeño and Baure varieties: the former has a root *-haka* while the latter has *-nuki/-noki*. Nevertheless, based on these Baure forms and on partial cognates in Mojeño and in Paunaka compounds meaning ‘beard, mustache’, as in Mojeño Trinitario *-hii-nuku* (Gill 1970: 7),¹ Jolkesky (2016: 19) reconstructs an etymon **-nuki* ‘mouth’ for his Proto-Mamoré-Guaporé (PMGU), the postulated common ancestor of Mojeño, Baure, Paunaka and Paikoneka (in section 5 I note potential cognates of this etymon in other branches of the Arawakan family).

In synthesis then, nothing certain can be said about the Terena noun *-pâho* ‘mouth’, which seems to remain historically unaccounted for and comparatively isolated within the family. First, it does not appear in any of the cognate sets on which Payne’s (1991) comparative study of the Arawakan family (and reconstruction of the Proto-Arawakan family) is based. Second, it is related only to the forms attested in Kinikinau by Matteson (1972) and by Walker & Ribeiro (2010), not a surprising or illuminating finding since Terena and Kinikinau are very close co-dialects. Finally, Terena *-pâho* ‘mouth’ bears no suggestive formal resemblance to the etymon reconstructed by Jolkesky (2016: 19), **-nuki* ‘mouth’, to the common ancestor of the Mojeño, Baure, Paunaka and Paikoneka. All of this could suggest that Terena *-pâho* ‘mouth’, despite its status as a basic vocabulary item (see e.g. Tadmor *et al.* 2010: 239) could be a loanword, perhaps, from a non-Arawakan language. As I show in the next sections, however, this form has a plausible Arawakan etymology linking it, at least, to its closest relatives, the Mojeño varieties.

4. Terena *pâho* ‘mouth’ and its cognates

The Terena noun *-pâho* [pa̯hɔ̯] ‘mouth’, is attested in virtually the same form from the earliest available documentation on the language: *<baho>* (Castelnau 1845, *apud* Martius 1867); *<bahó>* (Taunay 1868); *<pahotí>* (Schmidt 1903: 332). As stop consonant voicing is not distinctive in the language,² documented forms with initial ** are either too phonetic or reveal other

¹ As Jolkesky (2016) does not discuss Terena, I note that the closest match to this form I could find in Terena is *-inúku* ‘forehead’, though the semantics in this case suggests that the similarity is merely accidental.

² The Terena phonological segmental inventory is as follows, symbols having the standard IPA interpretations unless noted otherwise: *p, t, k, s, f, m, n, ŋ, w, j, r, h, a, e, i, o, u*. The mid vowels *e* and *o* are most frequently re-

inadequacies in the analysis. As discussed by Carvalho (2016a), some of the earliest records of Terena (Guaná) inalienable (dependent) nouns, such as body-part terms, actually present 1Psg possessive forms, a common form in elicitation, not the absolute or non-possessed forms implied by the accompanying glosses or translations.³ 1Psg is realized in Terena by a floating nasal feature that induces stop consonant voicing, in addition to the formation of a nasal-oral contour at the left edge of the leftmost obstruent consonant. Thus, given \emptyset -*pâho* 'his/her mouth', with \emptyset coding of 3P, one has m *bâho* 'my mouth' (see Eastlack 1968: 4). The circumflex diacritic \wedge used in written Terena indicates a falling pitch contour and a lengthened vowel. Finally, the form *<pahotî>*, recorded by Max Schmidt, includes the suffix *-ti* which indicates a generic or non-specific possessor (e.g. \emptyset -*hêwe* 'his/her foot', *hewêti* 'somebody's foot'; see e.g. Ekdahl & Butler 1979: 66 and next paragraph).

I have used a dash to indicate that *-pâho* like other inalienable nouns is a root. Hence, most body-part terms, many kinship terms and a few nouns denoting man-made objects or parts of wholes (e.g. 'root (of a tree)'), the semantic domains usually represented in the class of inalienable nouns, will always occur with some sort of morphological elaboration: in case they lack a prefix indicating the person-number features of a possessor (the prefix may be \emptyset -, as in the case of the Terena 3P),⁴ these items show up either with a suffix indicating their unpossessed status or in some other construction such as a nominal compound or incorporated within a verb. Suffixes signaling the unpossessed status of inalienable nouns are known as Absolute suffixes in the Arawakanist literature. In Terena the Absolute suffix has the form *-ti*, and is commonly glossed as expressing a generic, non-specific possessor. Thus: *-pâho* 'mouth' (a root, not a free-standing word), \emptyset -*pâho* 'his/her mouth', m *bâho* 'my mouth' but *pahotî* 'someone's mouth'. All other nouns, that is, alienable nouns such as *pirítaw* 'knife' or *wáteke* 'canoe', ordinarily occur as self-standing free forms in the absence of an explicit possessor. If marked for possession, however, additional morphology in the form of suffixes indicates their 'marked', possessed status: \emptyset -*pirítaw-na* 'his/knife', *wándéke-na* 'my canoe'. These suffixes are, in turn, called Genitive or Possessive suffixes in the Arawakanist tradition (see e.g. Payne 1990: 80–83; 1991: 378; Carvalho 2015).

A comparison of the candidates for cognate status in the meaning slot 'mouth' for the Bolivia-Paraná Arawakan languages reveals an obvious (that is, semantically-matched) cognate of the Terena form *-pâho* in Old Mojeño *<nupahò>* (Marbán 1702: 160). Also attested in Old Mojeño is an apparently competing form, *<nuhacà>*, which has a clear match in the form *-haka* 'mouth' attested in both modern Mojeño varieties, Ignaciano and Trinitario (see table 1). The latter is the Mojeño form presented in Jolkesky (2016: 13), no mention being made of Old Mojeño *<nupahò>*. The Old Mojeño forms contain the 1Psg possessive prefix *<nu->* and can

alized as open *ɛ* and *ɔ*, respectively, though I will adhere to the customary practice of using *e* and *o* in writing forms of the language. Note also that *w* is frequently realized as [v] or [β], though phonological patterns show it patterns like a sonorant, and that *j* is a marginally contrastive segment only.

³ Glossing conventions used in this paper are as follows: REFL: Reflexive; CONT: Continuative; TH: Thematic consonant, ACT: Actual Mood; AUX: Auxiliary verb; 3P: third person; 2P: second person; 1Psg: first person singular; 1Ppl: first person plural.

⁴ Though the person-number markers indicating possession in nouns or a subject in verbs are arguably prefixes (see e.g. Danielsen 2011: 508), some comment is necessary due to the non-concatenative exponence of some of these markers (Eastlack 1968: 4). As noted above, 1Psg is realized by a 'floating' nasal feature that spreads from the left edge of a word until blocked by an obstruent consonant, which becomes voiced. Thus, *éno* 'mother', *énō* 'my mother' but *ówoku* 'house', *ówôŋgu* 'my house'. The prefix *j-* indicates 2P in vowel-initial roots (cf. *j-ówoku* 'your house'), but fronting of the leftmost vowel in the word marks 2P elsewhere: \emptyset -*hána?**i-ti* 'he/she is big, tall', *hénâ?**i-ti* 'you are big, tall'. The status of these elements as prefixes is indicated by their left-aligned pattern of realization.

thus be reduced to the roots *<-pahò>* and *<-hacà>*. See the contrast with *<topahò cahacure>* ‘mouth of the river’, where *<-pahò>* occurs instead with the 3Psg non-human possessive *<to>* (Marbán 1702: 12, 160; cf. Rose 2015: 244).

Table 1. Forms for Mouth in the Bolivia-Paraná Arawakan Language

Language	Form	Source
Terena	<i>-pâho</i>	Author's field data
Old Mojeño	<i><nu-pahó></i> , <i><nu-hacà></i>	Marbán 1702: 160
Ignaciano	<i>-haka</i>	Ott & Ott 1983: 493
Trinitario	<i>-haka</i> <i>-nuku</i>	Gill 1970: 7 Françoise Rose, p.c.
Baure	<i>-noki</i>	Danielsen 2007: 469
Paunaka	<i>-niki</i>	Danielsen & Terhart 2014: 253

Both the Baure and Paunaka forms, in turn, show no obvious formal correspondence to Terena *-pâho* and to Mojeño *-haka*. As mentioned before, Jolkesky (2016: 19) considers the Baure and Paunaka forms cognates of a bound root *-nuku* found only in Mojeño forms such as *hii-nuku* ‘mustache’, all traced back to a PMGU etymon **-nuki* ‘mouth’. Recent data on the Trinitario variety shows, however, that *-nuku* does occur as a non-derived root in Mojeño, meaning either ‘mouth (of a person)’ or ‘neck of a bottle’ (Françoise Rose, personal communication).

The comparative data in table 1 suggests that Old Mojeño had a cognate of Terena *-pâho*, but that no such form was retained in the modern Mojeño varieties. Additional comparative data in table 2 below shows, however, that the relevant comparative patterns are more interesting than a simple case of vocabulary obsolescence in Ignaciano and Trinitario.

Table 2. Bolivia-Paraná forms for ‘door’

Language	Form	Source
Terena	<i>paha-péti</i>	Author's field data
Mojeño Ignaciano	<i>ta-paha</i>	Ott & Ott 1983: 344
Mojeño Trinitario	<i>ta-paho</i> , <i>-pahra</i>	Gill 1970: 33
Old Mojeño	<i><topaho></i>	Marbán 1702: 317
Baure	<i>haki-</i>	Danielsen 2007: 41
Paunaka	<i>nuineki</i>	Lena Terhart (p.c.)

The Terena form for ‘door’ is *pahapéti*, a nominal compound meaning ‘mouth of the house’ (cf. *péti* ‘house’; see e.g. Ekdahl & Butler 1979: 182). The structure of this form is in agreement with the general pattern for endocentric nominal compounds expressing part-whole relations or material provenance in Terena, which are regularly head-initial, as in *naūwaka* ‘beef’ (lit. meat-cow; *-naū* ‘meat, flesh’), *hēwetapi* ‘hen’s foot’ (foot-hen; *-hēwe* ‘foot, leg’). The earliest, published attestation of this form is *<pahapeti>* in Schmidt (1903: 566), who seems, however, to have overlooked the relation between this compound and the root for mouth (see Schmidt 1903: 593).

In relation to the other forms in table 2, note that the formative *ta-* in the Mojeño forms for ‘door’ is the 3P non-human possessor prefix (Ott & Ott 1983: 36; Rose 2015) also described as an ‘impersonal possessive’ (Gill 1970: 6). While Old Mojeño ‘mouth’ appears in Marbán (1702) with the 1Psg possessive prefix *<nu->*, the noun for ‘door’ has the same root preceded by the

3Psg non-human possessive *<to>* (see Rose 2015: 244). This is the same marker used with the other use of *<-paho>* denoting a non-human, metaphorically shifted sense, that of 'river mouth' noted above. In Trinitario one has a root *-paho* which, if possessed, shows the suffixation of the Possessive marker *-ra* and the effects of a recurrent process of vowel syncope (*paho-ra* > *pah-ra*; cf. Rose 2015: 253–254). The Paunaka form may be etymologizable to **nui-niki*, containing the root *-niki* ~ *-nîki* 'mouth', but this is not clear, and the remaining formative, *nui-*, is unattested elsewhere (Lena Terhart, personal communication).

Two facts in need of discussion and explanation are revealed by the data in table 2. First, that Terena *-pâho* 'mouth' has an allomorph *-paha* appearing, at least, in the compound *pahapéti* 'door'. Second, that Old Mojeño *<-pahò>* was polysemous, meaning either 'door' or 'mouth'. Note that this differs crucially from Terena *-pâho* 'mouth' and *pahapéti* 'door', two separate lexemes, even if clearly etymologically related. Accounting for the polysemy in the Old Mojeño form will be of vital importance for the diachronic developments postulated here but, first, I will deal briefly with the formal variation internal to Terena, that is, the *-pâho* ~ *-paha* allomorphy.

4.1. Terena **paho* > *paha*

Of the two forms for 'mouth' attested in Terena, *-pâho* and *-paha*, the former is the older, inherited (conservative) one, while *-paha* is innovative. Though the precise nature of the developments behind the emergence of the allomorph *-paha*, as well as their chronology, will remain an object for future investigation, pending a more thorough understanding of Terena historical phonology and morphology, the postulation of a single pre-Terena allomorph **-paho* 'mouth' seems to be plausible in view of the considerations below.

Formally, it is possible to derive *-paha* from *-pâho*, though not the reverse, by invoking a contextual factor such as vowel harmony or assimilation. The restriction of this process to forms such as the compound *pahapéti* 'door' may be explained on the basis of prosodic properties such as 'strength asymmetries' within the Foot or the Prosodic Word. In compounds such as *pahapéti*, from *pâho* 'mouth' and *péti* 'door', the first element usually loses its stress to the rightmost one (Bendor-Samuel 1961: 35), a rough indication of prosodic weakness. According to the evidence in Ekdahl & Butler (1979: 185), incorporated⁵ *-pahò* does retain its round vowel *o* where it is stressed, as a comparison of (2a) and (2b) clearly shows (the syllable bearing main stress is indicated in bold):⁶

(2) *Evidence for the role of stress placement in conditioning harmony*

- (a) Ø *-tímaru* *-paha* *-ʃ* *-o* *-wo*
 3P -lick -mouth-TH -ACT -REFL
 'He/she licked his own mouth'

⁵ Certain verbs in Terena, such as *kipó-* 'to wash', allow the incorporation of objects. Thus, given a root such as *-nône* 'face', one has *kipónonewoti* 'he/she washes his/her own face'. An alternative description, which seems to be favored in the literature, postulates the existence of a finite set of bound forms, called either 'qualifiers' (Ekdahl & Butler 1979: 185) or 'verbal classifiers' (Passer 2016) that can appear within the verbal word and may bear only an etymological relation to the independent noun roots.

⁶ As in other Arawakan languages, the morphemes glossed 'thematic' in Terena are affixes with little semantic content but which function as a kind of 'stem-closure' formative (Wise 1990: 90). In Terena these thematic suffixes are either *-ʃ* or *-k*. So-called athematic verb stems lack any thematic ending. Structures involving verbs and incorporated nouns function as stems which are either thematic or athematic, just like underived/simplex verb stems, hence the difference between the athematic stem in (2b) and the *-ʃ*-thematic stem in (2a).

- (b) *Ø -paru -pahó kójee*
 3P -open -mouth AUX
 'he/she is agape' (lit. 'he/she is with the mouth open')

Moreover, patterns that could be described as instances of *a* ~ *o* ablaut, often though not always associated with vowel harmony, are recurrent throughout the nominal and verbal morphology of Terena. In all cases, the overarching generalization is that *o* is basic and forms with *a* are derived. As noted in section 4, alienable nouns in Terena, as in other Arawakan languages, differ from inalienable nouns in bearing some additional morphological elaboration in possessive constructions, usually in the form of reflexes of one of the Possessive or Genitive suffixes reconstructed for the PA language. A subclass of these alienable nouns shows, when possessed, a change of every *o* to *a* (see Ekdahl & Butler 1979: 72, 182; a dash indicates a possessed form):

(3) Change of *o* to *a* in Terena alienable nouns

<i>sopôro</i>	<i>-sápara</i>	'maize'
<i>wôso</i>	<i>-wásá</i>	'line, thread'
<i>tôroro</i>	<i>-tárara</i>	'gourd'
<i>kohôfu</i>	<i>-kaháfa</i>	'oven'
<i>wojôre</i>	<i>-wájara</i>	'yam'

As the Possessive affixes in Arawakan languages are suffixal in nature, it is probable that the derivation of the possessed forms in (3) show the effects of vowel harmony targeting *o*. Elsewhere, change of *o* to *a* in what looks like a vowel harmony process is attested in Terena verbal morphology, where suffixation of the Irrealis (or Potential Mood) suffix *-a* triggers a change of every *o* in the verb stem to *a* (Ekdahl & Grimes 1964: 263; Ekdahl & Butler 1979: 46–47).

Whether the correct account of the emergence of the allomorph *-paha* will rely for the most part on morphological or prosodic considerations is unclear for now. However, the phonological and morphological patterns addressed above suggest that *-pahó* is the most conservative allomorph, *-paha* being derived by a recurrent process attested elsewhere in the language but conditioned by still unknown conditions in this specific case. This is enough for our present purposes.

4.2. Reconstruction of two etyma **paho* 'mouth' and **paho-peti* 'door': form and meaning

Based on the evidence displayed in tables 1 and 2, plus additional assumptions and facts that will be spelled out in detail below, I propose that a form essentially identical to **-paho* can be reconstructed to an earlier, common stage of development shared by Terena and Mojeño, for the meaning 'mouth' (see section 5 for discussion of what this common stage amounts to *vis-à-vis* the classification in (1)).⁷ A compound **-paho-peti* 'door' was also derived at this stage (see Ignaciano and Trinitario *peti* 'house'; Ott & Ott 1983: 300; Gill 1970: 34). It is plausible that after the Mojeño varieties innovated a separate lexeme to express the meaning 'mouth', *-haka*, also attested in Old Mojeño as a competing form along with a reflex of **-paho*, it was no longer

⁷ I say 'essentially identical' because the falling pitch contour and vowel length characteristic of Terena *-pahó* 'mouth', and of many other nouns and verbs in this language, has no accepted diachronic explanation at the moment. As these are clearly innovations of the language, I will deal only with the segmental content of the form, the one which is formally comparable to the cognates found in Mojeño.

necessary to employ the modifier *peti* to express the derived (or 'target') meaning 'door', hence the restricted meaning of *paho* as 'door' in the modern Mojeño varieties (*paha* in Ignaciano). As very few linguists and readers can be assumed to be familiar with the languages under discussion, and because there is little systematic historical investigation of the relations between Terena and Mojeño, it is perhaps fitting to present evidence that the reconstruction of **-paho* and **paho-peti* is supported by regular segmental correspondences, even if most of these turn out to be trivial identity correspondences (OM = Old Mojeño):

- (4) *Regular segmental correspondences supporting *-paho and *paho-peti.*
- (a) **p* > Ter. *p* : Ign. *p* : Trin. *p* : OM *p*
 FINGERNAIL Ter. *-hîpo* : Ign. *-hipaja* : Trin. *-hippo* : OM <*nuhipoñó*>; BONE Ter. *-ôpe* : Ign. *?iape* : Trin. *ópe-ra* : OM <*nuopè*>; DUCK Ter. <*pôhi*>, Trin. *pohi* : OM <*pohi*>; TO WASH Ter. *-kipo-* : Ign. *-sipaka* : Trin. *-sipko* : OM <*nusipocò*>; ROOT Ter. *pôe-hewe* : Ign. *ta-pare* : Trin. *-pore* : OM <*topore*>.
- (b) **a* > Ter. *a* : Ign. *a* : Trin. *a* : OM *a*
 HUSBAND Ter. *-îma* : Ign. *-ima* : Trin. *-ima* : OM <*niyma*>; GRANDSON Ter. *-ámori* : Ign. *-ámari* : Trin. *-amri* : OM <*nuamori*>; NAME Ter. *-îha* : Ign. *-ihare* : Trin. *-ihare* : OM <*nihà*>; TO HEAR Ter. *-kâmo* : Ign. *-sama* : Trin. *-samo* : OM <*nusamomoroicò*>; PERSON Ter. *âne* : Ign. *atfane* : Trin. *?fane* : OM <*achanè*>.
- (c) **h* > Ter. *h* : Ign. *h* : Trin. *h* : OM *h*
 NAME Ter. *-îha* : Ign. *-ihare* : Trin. *-ihare* : OM <*nihà*>; FINGERNAIL Ter. *-hîpo* : Ign. *-hipaja* : Trin. *-hippo* : OM <*nuhipoñó*>; TOUCAN Ter. *honô?e* : Ign. *hanare* : Trin. *hnore* ; TAIL Ter. *-îhi* : Ign. *tá-îhi-ki* : Trin. *-ihgi* : OM <*tahiqui*>; MOON Ter. *kohê?e* : Ign. *kahe* : Trin. *kóhe-ra* : OM <*cohè*>.
- (d) **o* > Ter. *o* : Ign. *a* : Trin. *o* : OM *o*
 BONE Ter. *-ôpe* : Ign. *?iape* : Trin. *ópe-ra* : OM <*nuopè*>; HOUSE (POSS.) Ter. *-pêno* : Ign. *-pena* : Trin. *-peno* : OM <*nupeno*>; BROTHER (ELDER) Ter. *-éfowi* : Ign. *-éfawî* : Trin. *-éfovi* : OM <*nechobi*>; MOON Ter. *kohê?e* : Ign. *kahe* : Trin. *kóhe-ra* : OM <*cohè*>; TOUCAN Ter. *honô?e* : Ign. *hanare* : Trin. *hnore* ; TO HEAR Ter. *-kâmo* : Ign. *-sama* : Trin. *-samo* : OM <*nusamomoroicò*>; EARTH, MUD Ter. *môte* : Ign. *mate-hi* : Trin. *mote-hi* : OM <*motehi*>; TO STEAL Ter. *-oméso* : Ign. *-ámetfa* : Trin. *-ómetfo* ; ROOT Ter. *pôe-hewe* : Ign. *ta-pare* : Trin. *-pore* : OM <*topore*>.
- (e) **e* > Ter. *e* : Ign. *e* : Trin. *e* : OM *e*
 Breast Ter. *-fêne* : Ign. *-tfene* : Trin. *-fene* : OM <*nuchene*>; BONE Ter. *-ôpe* : Ign. *?iape* : Trin. *ópe-ra* : OM <*nuopè*>; BROTHER (ELDER) Ter. *-éfowi* : Ign. *-éfawî* : Trin. *-éfovi* : OM <*nechobi*>; TOUCAN Ter. *honô?e* : Ign. *hanare* : Trin. *hnore* ; MOON Ter. *kohê?e* : Ign. *kahe* : Trin. *kóhe-ra* : OM <*cohè*>; MOTHER-IN-LAW Ter. *-imôse* : Ign. *-imase* : Trin. *-imse* : OM <*mose*>; TO STEAL Ter. *-oméso* : Ign. *-ámetfa* : Trin. *-ómetfo* ; ROOT Ter. *pôe-hewe* : Ign. *ta-pare* : Trin. *-pore* : OM <*topore*>.
- (f) **t* > Ter. *t* : Ign. *t* : Trin. *t* : OM *t*
 HEAD Ter. *-tûti* : Ign. *-futi* : Trin. *-futi* : OM <*nuchuti*>; BLOOD Ter. *iti* : Ign. *iti* : Trin. *iti* : OM <*iti*>; TERMITE Ter. *motôu* : Ign. *mata-ru* : Trin. *mto-ru* : OM <*motorù*>⁸; BROTHER (YOUNGER) Ter. *-âti*, Trin. *-ati*, OM <*nuati*>; NIGHT Ter. *jóti* : Ign. *jati* : Trin. *joti* : OM <*yati*>; BAT Ter.⁹ *witete* : Ign. *wite* : Trin. *vite* : OM <*vítê*>.

⁸ The gloss in Marbán (1702: 254) is 'hormigas, que comen la yuca recien plantada'.

⁹ Reduplication is a frequent property of animal names in Terena, as in *wâhaha* 'spider', *wétekeke* 'cayman', *fîrifîri* 'hummingbird' and *howâwo* 'frog'.

(g) *i > Ter. *i* : Ign. *i* : Trin. *i* : OM *i*

TAIL Ter. *-îhi* : Ign. *tá-ihi-ki* : Trin. *-ihgi* : OM <*tahiqui*>; HEAD Ter. *-tûti* : Ign. *-futi* : Trin. *-futi* : OM <*nuchuti*>; BLOOD Ter. *iti* : Ign. *iti* : Trin. *iti* : OM <*iti*>; BROTHER (YOUNGER) Ter. *-âti*, Trin. *-ati*, OM <*nuati*>; NAME Ter. *-îha* : Ign. *-ihare* : Trin. *-ihare* : OM <*nihâ*>; FINGERNAIL Ter. *-hîpo* : Ign. *-hipaja* : Trin. *-hippo* : OM <*nuhipoñó*>; TO WASH Ter. *-kipo-* : Ign. *-sipaka* : Trin. *-sipko* : OM <*nusipocò*>; JAGUAR Ter. *sîni* : Ign. *ifini* : Trin. *?fîni* : OM <*ichini*>; GRANDSON Ter. *-ámori* : Ign. *-ámari* : Trin. *-amri* : OM <*nuamori*>; BROTHER (ELDER) Ter. *-éfowi* : Ign. *-éfawî* : Trin. *-éfovi* : OM <*nechobi*>; NIGHT Ter. *jóti* : Ign. *jati* : Trin. *joti* : OM <*yati*>.

Though the reflex of *-paho was lost in modern Mojeño varieties as the expression of the meaning ‘mouth’, it was retained in the form for ‘door’ as a (slightly) obscure cognate, whose existence often points to the occurrence of semantic or functional shifts.

Correspondence (4d) above, directly relevant for the etymology Terena *-pâho*, Old Mojeño <*pahò*>, Trinitario *-paho*, Ignaciano *-paha*, requires further discussion. Given the lack of a phoneme *o* in Ignaciano (Ott & Ott 1959: 7–8; 1983: 5–7), this correspondence suggests that a merger **o*, **a* > *a* took place in this Mojeño variety. The operation of this merger was suggested by Rose (2015: 245, fn.3) and is explicitly advanced in Jolkesky’s (2016: 17) comparative work. Below I subject the relevant correspondence sets to scrutiny, bringing in the data from Terena which so far has not been included in the discussion.

In tables 3 and 4 I show cognate sets for two correspondences, one matching Terena *o* to Ignaciano *a* and Trinitario *o*, and the other having *a* in all three languages (corresponding vowels appear in bold).

Table 3. Correspondence set Ter *o* : Ign *a* : Trin *o*

	Terena	Ignaciano	Trinitario
Night	<i>jóti</i>	<i>jati</i>	<i>joti</i>
Rain, cloud	<i>ûko</i>	<i>uka</i>	<i>uko</i>
Earth, mud	<i>móte</i>	<i>mate</i>	<i>mote</i>
Grandmother	<i>-ôse</i>	<i>-atse</i>	<i>-otse</i>
Wife	<i>jêno</i>	<i>jena</i>	<i>jeno</i>
Shoulder, arm	<i>-pôwo</i>	<i>-pawa</i>	<i>-powo</i>
Tooth	<i>-ôe</i>	<i>-a?e</i>	<i>-o?e</i>

Table 4. Correspondence set Ter *a* : Ign *a* : Trin *a*

	Terena	Ignaciano	Trinitario
Sky	<i>wanúke</i>	<i>anu-ma</i>	<i>anu-mo</i>
Stone, stony floor	<i>marîpa</i>	<i>mari</i>	<i>mari</i>
Sun	<i>kátfe</i>	<i>satse</i>	<i>satse</i>
Person	<i>fâne</i>	<i>atfane</i>	<i>?tfane</i>
Son-in-law	<i>sîna</i>	<i>tjina</i>	<i>tjina</i>
Name	<i>-îha</i>	<i>-ihâ</i>	<i>-ihâ</i>
To hear	<i>-kâmo</i>	<i>-sama</i>	<i>-samo</i>

The distribution of the two correspondences — Ter *o* : Ign *a* : Trin *o* and Ter *a* : Ign *a* : Trin *a* — does not suggest any contextual factor that could point to a split in Terena and in Trinitario. Since the latter identity correspondence is non-controversially accounted by reconstructing **a*, a sensible assumption is to assign the former, non-identity correspondence to **o*, implying a merger of the two phonemes in Ignaciano.

Some complexities involving these correspondences should be noted. On the one hand, Old Mojeño data is equivocal, as it patterns with Trinitario in showing *o* matching Ignaciano *a* in most cases (see (4d)), but in many instances *a* is found instead in the Old Mojeño cognates of forms showing *o* in Trinitario and *a* in Ignaciano (e.g. OM *<tihapû>* ‘white’, Trin. *-hopu*, Ign. *-hapu*). My own intuition in this respect is that the Old Mojeño documents of Marbán are dialectally heterogeneous, not an implausible thesis in view of the multi-ethnic environment of the Christian Missions in 17th and early 18th century Llanos de Mojos. Sorting this problem out will demand a detailed investigation of Mojeño phonological diversification, a task beyond the immediate concerns of this paper. On the other hand, there are additional issues arising from attempts at establishing correspondences with languages elsewhere in Arawakan family and with the PA forms reconstructed by Payne (1991). Some of these issues were identified by Payne (1991: 472) himself. My own position, sketched in section 1 is that after Payne’s (1991) ambitious attempt at dealing with 24 distinct Arawakan languages in a single stroke, the time has come for comparative investigation of this family to proceed in a bottom-up manner, reconstructing from less inclusive intermediate subgroups. The correspondences amassed above, with the inclusion of data from Terena (arguably the closest relative of Mojeño within the family; see section 5 for discussion) robustly support the inference of a merger **o*, **a > a* in Ignaciano, a conclusion which, in my view, was already justified by comparing the known Mojeño varieties, notwithstanding the ambiguous testimony of Old Mojeño. I agree with one of the reviewers of this paper that the issue is not definitely closed, though I would add that bringing Terena data to the discussion not only tilted the balance in favor of the hypothesized merger of **a* and **o* in Ignaciano, but helped constitute a so far unacknowledged set of comparative patterns that must be successfully addressed by any competing explanation.

Before turning to questions of meaning, one should note that the postulation of a compound **paho-peti* ‘door’ (lit. ‘mouth (of the) house’) is consistent not only with the structure of endocentric, part-whole compounds in Terena, but also matches the structure of similar compounds in Mojeño. Compounds with modification structures, often involving two inalienable (hence, bound) lexemes follow the same order in this language. For the Ignaciano variety, for instance, Olza Zubiri *et al.* (2004: 219) note that for a root such as *-hija* ‘hair’ it is possible to derive *-hijatupa* ‘chest hair’ (cf. *-tupa* ‘chest’), *-hijasumu* ‘mustache’ (cf. *-sumu* ‘upper lip’), *-hijamama* ‘beard’ (cf. *-mama* ‘jaw, chin’), *-hija?a* ‘body hair’ (cf. *-?a* classifier for the body of humans or large animals).¹⁰

Reconstruction of the meaning ‘mouth’ for **-paho*, as opposed to the meaning ‘door’ attested in the modern Mojeño varieties, seems plausible on the grounds that (1) both Terena and Old Mojeño agree in this respect and (2) ‘mouth’ seems to be a more basic or salient mean-

¹⁰ The following comments are in order: I assume here a very simple notion of headedness for compounds, one based on meaning. As *paho-peti* ‘door’ is a ‘kind of opening’, *-paho* is taken to be the head of the construction. The same reasoning applies to Mojeño forms such as ‘mustache’ or ‘beard’, as all are distinct kinds of ‘hair’. Note also that the status of Mojeño *-hija?a* ‘body hair’ as a compound is debatable; one could treat classifiers synchronically along with more grammatical or functional markers and claim that *hija?a* is a suffixed noun. This is hardly problematic, however, as such bound classifiers are plausibly related, at least diachronically, to independent nominal lexemes.

ing than ‘door’ and, accordingly, metaphorical extensions relating these meanings are expected to operate in the direction ‘mouth’ > ‘door’, rather than in the reverse direction (with ‘mouth’ as the base and ‘door’ as the target). Similar semantic relations, such as ‘mouth’ > ‘estuary, mouth of river’ show a similar pattern where ‘mouth’ is the basic member of the relation (see Urban 2011: 12). In addition, the basic character of ‘mouth’ in relation to ‘door’ is reflected in the overt marking (in the sense of Urban 2011: 6) attested in Terena, where a nominal compound whose head is *paha* (< *paho*) is modified by the noun *peti* to express the meaning ‘door’ (literally “mouth of the house”), a pattern also attested elsewhere in the Arawak family (I return to this below).

Going beyond the mere classification of the semantic change from ‘mouth’ to ‘door’ as a metaphorical extension, standard assumptions make it likely that an intermediate stage of polysemy was involved in this shift (see e.g. Job 1982; Wilkins 1996; Urban 2011). On the model of Wilkins’ (1996: 269) graphic presentation of what he calls the ‘polysemous’ view of semantic change, the proposed relation between the Terena and Mojeño reflexes of the etymon *-*paho* can be depicted as follows:

(5) Stages:	(I)	(II)	(III)
Form:	*- <i>paho</i>	- <i>paho</i>	- <i>paho</i>
Meaning:	‘mouth’	‘mouth’ & ‘door’	‘door’

Moment (I) has a single form for ‘mouth’, even though, in a compound, it can be modified by the noun **peti* to express the notion ‘door’. This situation is what is attested for Terena, where both *-paho* ‘mouth’ and *pahapeti* ‘door’ co-exist. Stage (II) is characterized by the existence of polysemy, that is, the two related meanings ‘mouth’ and ‘door’ are associated with the sign *-paho*. This is the pattern attested in Old Mojeño: as seen in section 4, *-paho* can mean either ‘mouth’ or ‘door’, though the morphosyntactic context establishes one reading over the other. The Old Mojeño pattern provides the core link in a change relating ‘mouth’ and ‘door’ as it shows the existence of synchronic polysemy in a language that arguably represents an early, documented stage of languages that now have *-paho*, the form subject to change, only in the target meaning, in this case, ‘door’ (see Wilkins 1996: 269-270). Finally, stage (III) is that attested in the modern Mojeño varieties Ignaciano and Trinitario. The form *-paho* (-*paha* in the Ignaciano variety) is associated only with the meaning ‘door’, the meaning ‘mouth’ being now associated with a different lexeme *-haka*.

A reviewer suggests an alternative semantic reconstruction according to which a single polysemous form *-*paho* ‘mouth/door’ is posited. While Terena would have resolved the inherent ambiguity of the form by means of a compound *paho-peti* ‘door’, in Mojeño the meaning ‘mouth’ was taken over by the innovative form *-haka*, *-paho* being retained only for the meaning ‘door’. This scenario would be preferable for its greater simplicity, as it avoids the postulation of the compound **paho-peti* for any stage of the development of Mojeño, where this compound is, differently from Terena, unattested. Though I find the reviewer’s zeal against postulating this unattested compound structure for Mojeño history highly commendable, I will nevertheless stick to the view sketched above for two reasons: First, the polysemy associated with the Old Mojeño reflex of *-*paho* is still something in need of an explanation and, following Urban (2011: 24-29), it is plausible to think that the metaphorical extension behind the use of a form for ‘mouth’ as also meaning ‘door’ was formally mediated, in this case, by the compound **paho-peti*. That is, following Evans (2010) and Urban (2011) I take polysemy to provide a snapshot of semantic change in course, one that is mediated by the kind of formal structure such as nominal compounds.

Second, though it is true that **paho-peti* 'door' is not attested in Mojeño, external evidence in the form of recurrent compounds for 'door' derived from 'house' and 'mouth' in other languages/branches of the family make the postulation of **paho-peti* not entirely far-fetched. Examples include Yucuna *-numa* 'mouth', *numana* 'door' (Schauer *et al.* 2005: 231), Resígaro *pó?kónoomú* 'door', (Allin 1979: 442; Payne 1991: 408),¹¹ Bahuana *-numada* (Ramirez 1992: 121), Wapixana *panii-nom* 'door' (WLP 2000: 115), in these cases all having reflexes of a Proto-Arawakan root for 'mouth', reconstructed as **-numa* by Payne (1991: 413) (see also Ramirez 2001: 643). Moreover, the use of such compound expressions is independent of which Proto-Arawakan etymon for 'mouth' happens to be preserved in a language — remembering that Payne (1991) reconstructed three etyma for the meaning 'mouth' at the PA level. Languages like Paresi, which show reflexes of a different form, Payne's (1991: 413) PA etymon **k^hanaki* 'mouth', employ the same mechanism, with cognates of pan-Arawakan roots for 'house', as in *hati-kanatse* 'house-mouth', 'door' (see Brandão 2014: 248; Paresi *hati* 'house' is a cognate of Terena *péti*, **p>h* being a regular unconditioned development in Paresi).¹² It is therefore plausible to think that overt marking for the less basic meaning, in this case, 'door', was an intermediate stage in bringing about the polysemy seen in Mojeño (see Urban 2011: 25–29 for more general considerations).

In this section I have offered reasons to support the hypothesis that Terena *-pâho* 'mouth' has cognates in semantically-shifted modern Mojeño (Ignaciano and Trinitario) nouns for 'door', all being reflexes of earlier **-paho* 'mouth'. I have argued that this lexical semantic shift was formally mediated by a compound structure, **paho-peti*, expressing the meaning 'door', a pattern found throughout the Arawak language family. This is in agreement with the more general model of Urban (2011), relating overt marking with preferred directionality trends (as in 'mouth' > 'door') in diachronic semantics. Properties of the intermediate stages, including the existence of polysemy, are retained in the attested material on Old Mojeño, while Terena preserves the reconstructed overt-marking strategy for deriving the meaning 'door' as a compound involving the root for 'mouth' as the head element.

5. On the internal classification of Terena: The Achane branch hypothesis

In (1) I presented a working hypothesis on the internal classification of the Bolivia-Paraná languages, one that places Mojeño and Baure (along with other, less well-known languages such as Paunaka and Paikoneka) in one branch (the 'Bolivia' subgroup, presumably) and Terena as an independent, coordinate branch. As noted in section 2, this classification probably owes a lot to geographic factors and to non-conclusive assessments of 'relative linguistic proximity' (Walker & Ribeiro 2010: 3; Danielsen, Dunn & Muysken 2011: 185) — such as shared lexical retentions and structural similarities — that are consistent with but not indicative of subgroups. This classification has seemingly attained the status of orthodoxy, to the point that the best recent work on the historical-comparative linguistics of the Bolivia branch of the Bolivia-Paraná subgroup (Jolkesky 2016; labelled 'Mamoré-Guaporé') assumes the more distant position of Terena as a premise not worth discussing.

¹¹ Resígaro *-noomú* 'mouth' is a straightforward cognate of the other cited forms for 'mouth'. Resígaro has a single back rounded vowel, *o*, corresponding regularly to *u* in the other Northeastern languages such as Yucuna. Word-finally, **a>u* in Resígaro (Payne 1991: 473), where *<u>* stands for unrounded *u*.

¹² Though the use of 'mouth' as a base to express the target meaning 'door' is widespread among Arawakan languages, it is not a self-evident fact that it can be reconstructed at the Proto-Arawakan level. Some daughter languages, Wayuunaiki and Baniva de Maroa being two examples, use the root for 'eye' instead: in Wayuunaiki, given *-o?u* 'eye' and *piifí* 'house', one has *piifó?u* 'door' (see Captain & Captain 2005: 36). In Baniva de Maroa, *paníspuli* 'door' is a compound of *panísi* 'house' and *-puli* 'eye' (Mosonyi 2000: 504).

In the preceding sections of this paper I have relied heavily on Mojeño data to elucidate the etymology of a Terena ‘basic vocabulary’ noun, while Baure, the other (relatively) well-described member of the Bolivia-Paraná subgroup has contributed nothing. I was unable to find in this language any cognate of the Terena and Mojeño forms for ‘mouth’ or ‘door’ studied here, and this seems to agree with the comparative vocabulary amassed by Jolkesky (2016) in his study of the Bolivian languages.

I advance here the hypothesis (see (6) below) that Terena and Mojeño are more closely related to each other than any of these is to Baure, as an alternative to the scheme in (1).

(6) Achane branch and the internal classification of Bolivia-Parana Arawakan
Bolivia-Paraná subgroup

Baure
Paikoneka
Achane branch
Terena
Mojeño
Paunaka

To the best of my knowledge, this is the first time an internal classification of the Bolivia-Parana languages in which a branch composed of Terena, Mojeño and Paunaka, but excluding Baure and Paikoneka, is proposed (I will, for the moment, rely on Jolkesky 2016 for the assumption of a rather close relation between Baure and Paikoneka). The forms **-paho* ‘mouth’ and **paho-peti* ‘door’ can be taken as reconstructions for the common ancestor of this branch, or Proto-Achane (after the Old Mojeño noun for ‘person’). Moreover, the fact that Terena *-paho* and its previously undetected cognates in Mojeño have not been included in any compilation of Arawakan cognates (see section 3) is at least suggestive of its status as an innovation. Note that, in this respect, the etymon **-paho* differs from forms such as Baure *-noki* and Mojeño *-nuku*, given that the latter two, on the contrary, fit clearly within Payne’s (1991: 413) ‘Mouth3’ etymology, having cognates in Wayuunaiki *-aaniki* and Waurá *-kanati*, both meaning ‘mouth’ as well, and occurring in languages far apart from each other within the family. Therefore, an etymon close to Jolkesky’s (2016) **-nuki* ‘mouth’ can be plausibly assumed for Proto-Bolivia-Paraná, while **-paho* ‘mouth’ would constitute a shared innovation of the Achane languages. The best candidates I am aware of for the status of cognates of **-paho* ‘mouth’ outside of the Bolivia-Paraná subgroup are the forms attested in the languages of the Campa branch, such as Nanti and Matsigenka *-bagante* and Ashéninka *-paante* (see Michael 2011 and Heitzman 1973: 37). However, though Michael (2011) has successfully reconstructed the segmental phonology of the Proto-Campa language he offers no reconstructed etyma, there being no published reconstruction of the lexicon of Proto-Campa that supersedes the deeply flawed reconstruction of Matteson (1972). The present author is currently working on a lexical and morphological reconstruction of Proto-Campa and, if it turns out that the Proto-Campa etymon for ‘mouth’ is indeed a cognate of Proto-Achane **-paho*, this would invalidate its status as a Proto-Achane innovation.

Nevertheless, interesting additional evidence from the lexical and morphological domains furnish strong candidates for the status of shared innovations pointing to a stage of development common to Terena and Mojeño but not to Baure. The Baure root *-poʔe* ‘head’ (Danielsen 2007: 120) is used both as a syntactically independent noun (with appropriate morphology) or incorporated into a verb stem (Danielsen 2007: 126). Baure *-poʔe* is plausibly a cognate of Mojeño Ignaciano *-puʔi*, a classifier indicating round or spherical objects (Olza Zubiri *et al.* 2004: 286–288) and of Terena *-puʔi*, a classifier for ‘head-like’ objects (Ekdahl & Butler 1979:

167, 185). Moreover, this set probably reflects a much older etymon, as shown by apparent cognate forms attested in widely separate languages of the family, such as Baniva de Maroa *-bu* 'head' (Mosonyi 2000: 511), Yucuna *-pula* 'forehead' (Schauer & Schauer 2005: 205) and Garífuna *ábu-lugu* 'head' (Sabio & Ordoñez 2006: 7). Terena and Mojeño agree, however, in employing their cognates of Baure *-po?e* only as incorporated or bound classifiers; Terena *-tûti* and Mojeño *-futi* are used instead for 'head' when this is expressed as a syntactically independent expression. Terena *-tûti* and Mojeño *-futi* are, like Proto-Achane **-paho*, reasonably good candidates for being innovations. Payne (1991: 405) included *-tûti* and *-futi* in his cognate set for PA 'forehead' though this is, for diverse reasons, a questionable etymology. Note, first, that the final syllables of both forms are arbitrarily excised from the comparison; there are, however, no clear grounds for analyzing these as *-tu-ti* and *-futu-ti*, respectively, at any level. Second, Payne proposes that Ashéninka and Matsigenka reflexes of Proto-Campa **gi-to* 'head' (Matteson 1972: 213) present, in the final syllable *-to*, a cognate of the *-tu-/futu-* formative he identifies in the Terena and Mojeño forms. Recent and more extensive documentation of Ashéninka varieties reveals, however, that Payne's (1991: 405) analysis of *-to* in **gi-to* as a classifier meaning 'head-shaped, round' is incorrect; *-to* is, indeed, a classifier, but its meaning is 'hollow, long, rigid' (see Mihas 2015: 414), where the meaning 'rigid' is probably the one relevant for 'head' (see that the remaining morpheme **-gi-* is a straightforward reflex of the PA etymon **kiwi* 'head' reconstructed by Payne 1991: 407).

Certain specific morphological patterns attested in both Terena and Mojeño, but not in Baure, also imply a rather close structural similarity between the putative members of the Achane branch, and preliminary inspection of comparative data suggest that these could be shared innovations. A 1Ppl verbal suffix is usually reconstructed as **-wa* or **-w(a)* at the PA level or at another intermediate level (see Aikhengvald 1999: 88; Danielsen 2011: 514–515, the latter for 'Proto-Southern-Arawakan'). In Mojeño, this suffix has not only a final *i* vowel that seems characteristic of some southern Arawakan languages (see Danielsen 2011: 215) but differs as well in having a -VCV structure, appearing as *-avi* (Rose 2015: 244). Interestingly, this initial vowel of the Mojeño 1Ppl suffix, unattested anywhere else in the family, furnishes an explanation for a morphophonological quirk of Terena: in this language, the 1Ppl has the form *-wi* but it is unique among all person-marking suffixes in the language in that it triggers the lengthening of a preceding vowel (see Eastlack 1968: 5; Ekdahl & Butler 1979: 35). Thus, contrasting with *perésa-nu úne* '(you) give me water!', with the 1Psg object suffix *-nu*, one has *perésaa-wi úne* '(you) give us water!', with the 1Ppl object suffix *-wi* triggering lengthening of the final vowel of the verb stem (data from Ekdahl & Butler 1979: 33). This lengthening effect on a preceding vowel can be easily explained as the result of *sandhi* processes (compensatory lengthening) involving absorption of the vowel present in the Mojeño cognate suffix, a suffix-initial vowel not reconstructed for the PA language.

There is, I submit, enough reasons to consider the existence of a branch I label Achane, including Terena and Mojeño but excluding Baure, as a credible alternative to the geographically-based scheme in (1). At this point, however, the existence of this Achane branch is no more than a hypothesis worth investigating; it goes without saying that further investigation of potential shared innovations in lexicon, morphology and phonology is necessary before any definite conclusions can be attained.

6. Conclusion and final remarks

This paper demonstrated, with material from a language that is highly understudied from a diachronic standpoint, how etymological analysis, if properly conceived and conducted, can

considerably increase our understanding of the historical development of a language and its relatives.

I have shown here that Terena *-pâho* ‘mouth’ has cognates in the different speech varieties forming the Mojeño language. Accounting for this etymology calls, however, for an explicit hypothesis relating *-pâho* ‘mouth’ to forms in Ignaciano and Trinitario whose meaning is ‘door’, thus implying the action of a lexical semantic shift. Evidence from Old Mojeño was crucial in that it shows a stage in which a reflex of **-paho* is associated with synchronic polysemy, which is predicted to exist given our general understanding of how semantic change proceeds (see Wilkins 1996: 269–270; Urban 2011). The whole account proposed involves the postulation of two proto-forms, a root **-paho* ‘mouth’ and a nominal compound **paho-peti* ‘door’, derived with the use of **peti* ‘house’, as a modifier of the noun **-paho*. It is plausible that after the Mojeño varieties innovated a separate lexeme to express the meaning ‘mouth’, *-haka*, also attested in Old Mojeño as a competing form along with a reflex of **-paho*, it was no longer necessary to employ the modifier *peti* to express the derived (or ‘target’) meaning ‘door’, and, consequently, reflexes of **-paho* came to mean ‘door’ exclusively in these speech varieties. The etymological account proposed is consistent not only with the usual formal, that is, phonological and morphological constraints on compelling etymologies, but is also consistent with constraints on semantic reconstruction. Reference to the more general and widespread character of formations for ‘door’ as a compound involving nouns for ‘mouth’ and ‘house’, in particular to the ubiquity of this pattern in the Arawakan language family, also meets one of the demands usually placed on credible semantic developments (see Job 1982). In the end, a rather close relationship between Terena and Mojeño is suggested by the argumentation presented here, advancing the hypothesis of a branch composed of the most recent and exclusive common ancestor of these two languages as a viable proposal for internal classification.

References

- Aikhenvald, Alexandra. 1999. The Arawak Language Family. In: R. M. W. Dixon, Alexandra Aikhenvald (eds.) *The Amazonian Languages*. Cambridge University Press: 65–106.
- Brandão, Ana Paula. 2014. A Reference Grammar of Paresi-Haliti (Arawakan). Ph.D. Dissertation, The University of Texas at Austin.
- Campbell, Lyle. 2012. Classification of the Indigenous Languages of South America. In: Lyle Campbell, Verónica Grondona (eds.) *The Indigenous Languages of South America: A Comprehensive Guide*. De Gruyter Mouton: 59–166.
- Captain, David, Linda Captain. 2005. *Diccionario Básico Ilustrado Wayuunaiki- Español, Español-Wayuunaiki*. Bogotá: Editorial Buena Semilla.
- Carvalho, Fernando O. de. 2015. On the Realization of Nominal Possession in Mehinaku: A Diachronic Account. *International Journal of American Linguistics* 81 (1): 119–132.
- Carvalho, Fernando O. de. 2016a. Terena, Chané, Guaná and Kinikinau are one and the same language: Setting the record straight on Southern Arawakan linguistic diversity. *LIAMES* 16 (1): 39–57.
- Carvalho, Fernando O. de. 2016b. Internal and Comparative Reconstruction in Yawalapiti: Palatalization and Rule Telescoping. *International Journal of American Linguistics* 82 (3): 285–316.
- Carvalho, Fernando O. de. 2016c. The Diachrony of Person-Number Marking in the Lokono-Wayuunaiki Subgroup of the Arawakan Language Family: Reconstruction, Sound Change and Analogy. *Language Sciences* 55: 1–15.
- Carvalho, Fernando O. de. Forthcoming. Fricative Debuccalization and Primary Split in Terena (Arawakan) Historical Phonology. *International Journal of American Linguistics* (July/2017 issue).
- Danielsen, Swintha. 2007. *Baure: An Arawakan Language of Bolivia*. Indigenous Languages of Latin America (ILLA). University of Leiden.
- Danielsen, Swintha. 2011. The Personal Paradigms in Baure and other Southern Arawakan Languages. *International Journal of American Linguistics* 77 (4): 495–520.

- Danielsen, Swintha, Lena Terhart. 2014. Paunaka. In: Mily Crevels, Pieter Muysken (eds.) *Lenguas de Bolivia, Vol. III: Oriente*. La Paz: Plural Editores: 221–258.
- Danielsen, Swintha, Michael Dunn, Pieter Muysken. 2011. The Spread of the Arawakan Languages: A View from Structural Phylogenetics. In: Alf Hornborg, Jonathan D. Hill (eds.) *Ethnicity in Ancient Amazonia*. University Press of Colorado: 173–195.
- Eastlack, Charles. 1968. Terena (Arawakan) Pronouns. *International Journal of American Linguistics* 34 (1): 1–8.
- Ekdahl, Elizabeth, Nancy Butler. 1969. Terêna Dictionary. Arquivo Lingüístico 095. Brasília: Summer Institute of Linguistics.
- Ekdahl, Elizabeth, Nancy Butler. 1979. *Aprenda Terêna*. Brasília: Summer Institute of Linguistics (SIL).
- Evans, Nicholas. 2010. Semantic Typology. In: Jae Jong Song (ed.) *The Oxford Handbook of Linguistic Typology*. Oxford: Oxford University Press: 504–533.
- Gill, Wayne. 1957. *Conversación y Gramática Trinitaria*. San Lorenzo de Mojos: Misión Evangélica Nuevas Tribus.
- Gill, Wayne. 1970. *Diccionario Trinitario-Castellano y Castellano-Trinitario*. Misión Evangélica Nuevas Tribus.
- Heitzman, Allene. 1973. Reflexes of Some Proto-Campa Consonants in Modern Campan Languages and Dialects. MA Thesis, University of Kansas.
- Hill, Jonathan D., Fernando Santos-Granero. 2002. *Comparative Arawakan Histories*. Chicago: University of Illinois Press.
- Hornborg, Alf, Jonathan D. Hill. 2011. *Ethnicity in Ancient Amazonia*. Boulder: University Press of Colorado.
- Job, D. M. 1982. Semantic Change and Etymologies. In: Anders Ahlqvist (ed.) *Papers from the 5th International Conference on Historical Linguistics*. Amsterdam: John Benjamins: 163–170.
- Jolkesky, Marcelo. 2016. Uma Reconstrução do Proto-Mamoré-Guaporé (Família Arawakan). *LIAMES* 16 (1): 7–37.
- Kaufman, Terrence. 1990. Language History in South America: What We Know and How to Know More. In: Doris Payne (ed.) *Amazonian Linguistics: Studies in Lowland South American Languages*. Austin: University of Texas Press: 13–67.
- Kaufman, Terrence. 1994. The Native Languages of South America. In: Christopher Moseley, R. E. Ascher (eds.) *Atlas of the World's Languages*. Routledge: 46–76.
- Lawrence, Aimee. 2014. Reconstruction of Proto-Kampa Verbal Morphology. *Proceedings of the Annual Meetings of the Berkeley Linguistics Society 2014*: 256–271.
- Mailhammer, Robert. 2014. Etymology. In: Claire Bowern, Bethwyn Evans (eds.) *The Routledge Handbook of Historical Linguistics*. London and New York: Routledge: 423–441.
- Marbán, Pedro. 1702. *Arte de la Lengua Moxa*. Lima: Imprenta Real de Joseph de Contreras.
- Martius, Karl Friedrich von. 1867. *Beiträge zur Ethnographie und Sprachenkunde Amerikas zumal Brasiliens*. Leipzig: Friedrich Fleischer.
- Matteson, Esther. 1972. Proto-Arawakan. In: Esther Matteson (ed.) *Comparative Studies in Amerindian Languages*. The Hague: Mouton.
- Michael, Lev. 2009. Review of Ramirez 2001. *International Journal of American Linguistics* 75 (3): 447–450.
- Michael, Lev. 2011. *La Reconstrucción y la Clasificación de la Rama Kampa de la Familia Arawakan*. Presented at CILLA V, 2011.
- Mihas, Elena. 2015. *A Grammar of Alto Perené (Arawak)*. Berlin: De Gruyter Mouton.
- Mosonyi, Esteban E. 2000. Introducción al Análisis del Idioma Baniva. In: María S. González de Pérez, María L. Rodríguez de Montes (eds.) *Lenguas Indígenas de Colombia: Una Visión Descriptiva*. Santafé de Bogotá: Instituto Caro y Cuervo: 499–513.
- Olza Zubiri, Jesús, Conchita Nuni de Chapi, Juan Tube. 2004. *Gramática Moja Ignaciana (Morfosintaxis)*. San Cristóbal: Universidad Católica de Táchira.
- Ott, Willis, Rebecca B. de Ott. 1959. Fonemas de la Lengua Ignaciana. *Notas Lingüísticas de Bolivia*, 1. La Paz: Instituto Lingüístico de Verano (ILV): 1–13.
- Ott Willis, Rebecca B. de Ott. 1983. Diccionario Ignaciano y Castellano Con Apuntes Gramaticales. Cochabamba: Summer Institute of Linguistics (SIL).
- Passer, Matthias B. 2016. (What) Do Verbal Classifiers Classify? *Lingua* 174: 16–44.
- Payne, David. 1990. Some Widespread Grammatical Forms in South American Languages. In: Doris Payne (ed.) *Amazonian Linguistics: Studies in Lowland South American Languages*. Austin: University of Texas Press: 75–87.
- Payne, David. 1991. A Classification of Maipuran (Arawakan) Languages Based on Shared Lexical Retentions. In: Desmond Derbyshire, Geoffrey K. Pullum (orgs.) *Handbook of Amazonian Languages*. Vol. 3. Mouton de Gruyter: 355–499.

- Ramirez, Henri. 1992. *Le Bahuana: Une Nouvelle Langue de la Famille Arawak*. Amerindia 17.
- Ramirez, Henri. 2001. *Línguas Arawakan da Amazônia Setentrional*. Manaus: Editora da Universidade do Amazonas.
- Rose, Françoise. 2015. Innovative Complexity in the Pronominal Paradigm of Mojeño: A Result of Contact? F. Gardani *et al.* (eds.) *Borrowed Morphology*. Berlin: The Gruyter Mouton: 239–265.
- Sabio, Fernando, Celia K. Ordoñez. 2006. *Hererun Wagüchagu*. La Ceiba: Asociación Misionera Garífuna de Honduras.
- Schmidt, Max. 1903. Guaná. *Zeitschrift für Ethnologie* 35 (2/3): 324–336.
- Tadmor, Uri, Martin Haspelmath, Bradley Taylor. 2010. Borrowability and the Notion of Basic Vocabulary. *Diachronica* 27 (2): 226–246.
- Taunay, Alfredo. 1868. *Scenas de Viagem: Exploração entre os Rios Taquary e Aquidauna no Distrito de Miranda*. Rio de Janeiro: Typographia Americana.
- Urban, Matthias. 2011. Asymmetries in Overt Marking and Directionality in Semantic Change. *Journal of Historical Linguistics* 1 (1): 3–47.
- Walker, Robert, Lincoln A. Ribeiro. 2010. Bayesian Phylogeography of the Arawakan Expansion in Lowland South America. *Proc. R. Soc. B* 278: 2562–2567.
- Wilkins, David. 1996. Natural Tendencies of Semantic Change and the Search for Cognates. In: Mark Durie, Malcolm Ross (eds.) *The Comparative Method Reviewed: Regularity and Irregularity in Language Change*. Oxford University Press: 264–304.
- Wise, Mary Ruth. 1990. Valence-Changing Affixes in Maipuran Arawakan Languages. In: Doris Payne (ed.) *Amazonian Linguistics: Studies in Lowland South American Languages*. Austin: University of Texas Press: 89–116.
- WLP (Wapixana Language Project). 2000. *Scholar's Dictionary and Grammar of the Wapishana Language*. Lethem: Wapishana Language Project and SIL International.

Фернанду де Карвалью. Этимология терена (аравакской) основы 'рот' и ее значение для внутренней классификации аравакских языков

В статье рассматривается этимология существительного *-pâho* 'рот' в языке терена. В современном аравакском языкоznании принято считать, что в других языках этой семьи у нее нет когнаторов; автор, однако, показывает, что такие когнаторы все же есть в языке мохеньо, близкородственном терена (оба входят в боливийско-паранскую подгруппу аравакской семьи). В игнасиано и тринитарио (два наиболее хорошо изученных диалекта мохеньо) эти когнаторы подверглись семантическому сдвигу и стали означать 'дверь'. Автор описывает семантические и формальные связи между этими формами через реконструкцию простой основы **-paho* 'рот' и композита **paho-peti* 'дверь' (букв. 'рот /дома/'). Данная этимологизация имеет значимость и с точки зрения внутренней классификации аравакских языков: автор предполагает, что терена и мохеньо составляют отдельную ветвь аравакской семьи — подгруппу ачане (в которую не входит язык бауре).

Ключевые слова: аравакские языки, язык терена, семантические сдвиги