NIMUENDAJÚ WAS RIGHT: THE INCLUSION OF THE JABUTÍ LANGUAGE FAMILY IN THE MACRO-ĴETO STOCK

EDUARDO RIVAIL RIBEIRO
MUSEU ANTROPOLÓGICO, GOIÂNIA
UNIVERSITY OF CHICAGO

HEIN VAN DER VOORT
RADBOUT UNIVERSITEIT NIJMEGEN
MUSEU PARAENSE EMÍLIO GOELDÍ

So far, the available data on the Arikápú and Djeoromitxí languages of the Brazilian Amazon, which together form the Jabutí language family, have been too sparse to enable one to evaluate existing proposals for a genealogical classification. Recent descriptive fieldwork, however, has yielded information that strongly corroborates an early hypothesis of Nimuendajú (2000 [1935]) that the languages are related to those of the Jê family. In this article, we compare the reconstructed ancestors of both families, Proto-Jabutí (van der Voort 2007) and Proto-Ĵê (Ribeiro [forthcoming]), ultimately demonstrating that Nimuendajú was right. We also discuss some of the consequences of this discovery for

1 The Arikápú and Djeoromitxí data in this article are from the fieldwork of Hein van der Voort and the data on other Macro-Ĵê languages are from the fieldwork of Eduardo Ribeiro, unless indicated otherwise. We want to thank our teachers and consultants Nazaré Arikápú, the late Mamoa Arikápú, Raimundo Jabutí, André, Armando, and José Roberto Jabutí, and José Puriété Arikaprú (Arikápú and Djeoromitxí); the late Ijeseberi Karajá, Sinvaldo Wahuka, and Luíís Kurikala (Karajá); and Ataíde Xehitãã, Marilda Xahtã, and José Koy (Ofayé). We are gratefully indebted to Denny Moore for his encouragement and ideas. We also gratefully acknowledge the beautiful map tailor made by Willem Doelman. We thank the University of Oregon at Eugene, the Universidade Federal do Pará at Belém, and the Netherlands Organisation for Scientific Research (NWO), who provided infrastructure and opportunities for collaboration through NWO internationalization grant no. 235-70-002 “Exploring the Linguistic Past: Historical Linguistics in South America.” Eduardo Ribeiro would like to thank the members of the research group “Estudos Histórico-Comparativos Macro-Ĵê” at the Museu Antropológico (Universidade Federal de Goiás), particularly Christianne de Oliveira and Marília Ferreira, for sharing their knowledge of Apinajé and Parkatêjê, respectively. We also want to thank Spike Gildea, José Pedro Viegas Barros, and two anonymous reviewers for their constructive comments. Ribeiro is furthermore especially thankful to Luciana Dourado, for providing him with data of (and clarifications on) Panará, and to Lucy Seki, for valuable insights on Krenák. Ribeiro’s research on Ofayé was undertaken under the auspices of the Endangered Languages Documentation Project (SOAS, London), whose support he gratefully acknowledges. Hein van der Voort wants to thank Sebastian Drude, Vilacy Galucio, Gale Goodwin Gómez, and Gessiane Picanço for their important comments and Sérgio Meira for teaching him about historical reconstruction. Furthermore, he gratefully acknowledges the support of the Netherlands Foundation for the Advancement of Tropical Research (WOTRO), who financed the entire Arikápú descriptive project under grant no. W 39-273, and the Netherlands Organisation for Scientific Research (NWO), who financed the Jabutí comparative research grant no. 276-70-005. Notwithstanding our indebtedness, the responsibility for the contents of this article rests solely with us.
current ideas about prehistoric population movements in lowland South America.

[KEYWORDS: Arikapú, Djeoromitxí, Jabutí family, Macro-Jê stock, Amazonian languages]

1. Introduction. Arikapú and Djeoromitxí are two Brazilian Amazonian languages that together compose the Jabutí language family. Both languages are traditionally spoken on the Rio Branco, a right tributary of the Guaporé River that separates the northeastern Bolivian province of Bení from the western Brazilian state of Rondônia. Although these languages used to be native to traditional rainforest populations of considerable size until about a century ago, the (estimated) number of Djeoromitxí speakers has now dwindled to a mere 50, and Arikapú is on the verge of extinction with only one elderly speaker. The total number of persons that identify ethnically with Arikapú or Djeoromitxí is somewhat bigger.

Until recently, the Jabutí languages were practically undocumented. Nevertheless, claims have been raised about their classification. On the basis of 100 words, Rivet (1953) classified one of the Jabutí languages as a Chibcha language. Loukotka (1963; 1968) grouped the Jabutí languages together as a small isolated family on the basis of short comparative word lists. Nimuendajú (2000 [1935]) and Greenberg (1987) classified the languages as belonging to the Macro-Jê stock on the basis of even smaller word lists.2

In this article we provide sound evidence, on the basis of extensive and reliable new data from all language components, for the classification of the Jabutí languages as belonging to the Macro-Jê language stock. To this end, we first present the general characteristics of both families and their respective proto-languages (2 and 3). Then we show correspondences between Proto-Jabutí and Proto-Jê (4). Finally, we discuss some possible consequences of such a remote westerly location of a Macro-Jê family for current ideas about the original dispersion center of the Macro-Jê-speaking tribes. The Swadesh basic 100–word list for Arikapú and Djeoromitxí is provided in Appendix A, together with reconstructed proto-forms.

---

2 As those who are familiar with South American linguistic historiography will know, Nimuendajú was careful not to make claims of linguistic connections (genetic or otherwise) without enough evidence (as illustrated by his debate with Mansur Guérios, published as Nimuendajú and Guérios 1948). The evidence he mentions, included in a personal letter (and not as a paper for publication), is indeed very little, but it is likely that he had detected many more similarities. His comparison includes items from the Kamakã and Kotoxó languages (both belonging to the Kamakã family, which had traditionally been considered as closely related to Jê since von Martius’s [1867] classification), in addition to languages now included in the Jê family. The fact that Nimuendajú, in his latest classification (IBGE 1981 [1944]), excludes from Jê all the traditional eastern Macro-Jê families (such as Kamakã) while maintaining the Jabutí languages is in itself symptomatic of his degree of certainty that this family should be included.
2. The Jabutí languages. In this section, we present a sketch of the sociolinguistic (2.1) and historical (2.2) contexts of Arikapú and Djeoromitxí and give a basic typological, phonological, and grammatical characterization of the languages (2.3) and their ancestor Proto-Jabutí (2.4), as reconstructed by van der Voort (2007).

2.1. Sociolinguistic setting. Traditionally, the Jabutí languages were spoken in close proximity to one another, to the Tuparí languages (Tupí family), and possibly also to Chapakura languages and isolated (or unclassified) languages. The steep demographic decline after the onset of regular contact with Western society has been characteristic of most native peoples in the Guaporé region. Today, Rondônia represents one of the most deforested regions of the Amazon and the decimated indigenous groups live scattered over different indigenous reserves, agricultural boomtowns, and city slums. The great majority of its 25 indigenous languages have fewer than 50 speakers.

Nowadays, the majority of the Djeoromitxí and one speaker of Arikapú live in the Guaporé indigenous reserve. This reserve, formerly called Ricardo Franco, was established in the 1930s to harbor the remnants of indigenous groups of southern Rondônia—individuals or groups that had escaped from slavery on rubber plantations and other Western extractivist enterprises or that had become dispossessed by invasions or by the effects of governmental policies of land distribution. The reserve is located on the Guaporé River itself and is only accessible by boat or by aircraft. Although the river represents the border with Bolivia and the indigenous inhabitants could not have been pushed off the map of Brazil any further, its remote location provides at least some chance for cultural survival. Presently, approximately 500 individuals from ten different ethnic backgrounds inhabit the reserve. Portuguese is the main lingua franca. The minority languages Djeoromitxí and Makuráp (Tuparí branch of Tupí) are spoken by all generations. There is also a family of Tuparí (Tuparí) speakers and there are several, mostly elder, speakers of Aruá (Mondé branch of Tupí), Wayurú (Tuparí), Kuyubí and Wari’ (both Chapakura), and Aikanã and Kanoê (both isolates) The ethnic Arikapú either have Djeoromitxí or Makuráp as a first language.

Another speaker of Arikapú passed away in 2009. He and the remaining speakers of Djeoromitxí lived closer to their traditional lands in the Rio Branco indigenous reserve, which was demarcated in the 1980s and covers part of the traditional habitat of the Makuráp and Djeoromitxí. As a reserve, this area is governmental property and by the Brazilian constitution only the indigenous inhabitants enjoy usufructuary rights of the forest and the rivers that run through it, of which the Rio Branco is the principal one. Unfortunately, the headwaters of the Rio Branco, which form the traditional habitat of the Djeoromitxí, Arikapú, and Tuparí, were left outside of the demarcated area. Every year more and more of the region is deforested and the rivers are
blocked or diverted for hydroelectric plants. This damages the hydrological and ichthyologic balance of the reserve and contributes to rising levels of agro-toxins.\(^3\) The integrity of the reserve is further threatened by invading Western loggers, poachers, and illegal fishermen. At this moment, approximately 400 persons from six different ethnic backgrounds inhabit the reserve. Tupari (Tuparí) is the dominant language in most settlements, while Makuráp (Tuparí) is spoken by a minority. Djeoromitxí and Aruá (Mondé) are spoken by several, mostly elderly, persons. There is a group that identifies itself as Kampé but does not speak the traditional language (which seems to be a Tuparí language). The ethnic Arikapú who live in the Rio Branco reserve all have Tuparí as a first language. Many inhabitants of the reserve also speak Portuguese.

Even though the number of speakers of Djeoromitxí is low, it is a living language in several settlements within the Guaporé reserve. The Arikapú language went out of use years ago and may show phenomena of obsolescence.

2.2. History of research. The first documentation ever of a Jabutí language is the word list from 1914 by the English colonel Percy Harrison Fawcett, who in the early twentieth century led several expeditions to survey the Bolivian–Brazilian frontier. Fawcett recorded a list of about 100 words on one of the eastern tributaries of the upper Rio Branco and refers to the speakers of the language as Mashubi (1915; 1953; see also Rivet 1953).\(^4\) Fawcett estimated that the population numbered about 2,000 individuals, divided over 24 villages. Since Fawcett, no one has ever again met a group called Mashubi. Therefore, the Mashubi language is considered in most of the literature (Loukotka 1963; 1968, Campbell 1997, and Dixon and Aikhenvald 1999) as an extinct language, related to Arikapú and Djeoromitxí. However, the Swiss ethnographer Franz Caspar, who lived with the Tupari for many months in 1948 and 1954–55, and who also documented the languages of their Arikapú and Djeoromitxí neighbors, claims that Mashubi is probably a variety of Arikapú (Caspar 1955a; 1955b). This is confirmed by evidence from the field (van der Voort 2005; forthcoming a). Hence, the Jabutí languages are considered in this article as forming a small family with only two

\(^3\) In addition, the construction of hydroelectric dams are a danger to archaeological remains that represent evidence of traditional indigenous habitation. At the construction site of a nearby new house in 2004, funerary urns were unearthed and destroyed. Among these urns were possibly those in which our consultant, Nazaré Arikapú, had buried her mother and her first child. Brazilian law considers indigenous funerary remains as evidence in matters of land rights. No decisive action was taken by the authorities, however.

\(^4\) During this trip, Fawcett met the Swedish ethnographer Erland Nordenskiöld, who was heading for the Huari (Aikanã) (Nordenskiöld 1915). Somehow Nordenskiöld obtained a copy of Fawcett’s word list and eventually passed it on to the French linguist Paul Rivet, who published it in 1953, trying to relate the language to Chibcha.
members: Arikapú and Djeoromitxí; and the earliest documentation of Arikapú is that by Fawcett (in Rivet 1953).

The next phase in documentation occurred when the German ethnographer Emil Heinrich Snethlage visited the Guaporé region in 1933–34 (Snethlage 1937). Snethlage met many peoples of the Rio Branco, including the Arikapú and Djeoromitxí (whom he also called Kipiu), and returned to Berlin with a collection of objects, photographs, film, and music recordings on wax rolls (Snethlage 1939). Unfortunately, his early death (see R. Snethlage 2002) prevented him from publishing about his experiences and data to a fuller extent. Snethlage had been in contact with the famous Curt Nimuendajú, who during his life visited over a 100 indigenous peoples and published widely on their languages, cultures, and mythology during the first half of the twentieth century. Snethlage provided him with short word lists of Arikapú and Djeoromitxí. Nimuendajú noticed a striking resemblance to words in Jê languages such as Xerénte, Kayapó, Kaingáng, and Timbíra. He hypothesized that both languages compose a branch of the Jê language family, which he mentioned in 1935 in a letter to his colleague Carlos Estevão de Oliveira (Nimuendajú 2000:219–22). Also, he represented them as such on a manuscript map (Nimuendajú 2000:10, figure 7), and on his classificatory map, which was eventually published in 1981 (IBGE 1981). The very fact that Greenberg (1987) adopted this classification may have been a reason that it did not become broadly accepted among Americanists.

As mentioned above, Franz Caspar lived in the Rio Branco region between 1948 and 1955 and is known especially for his popular and scientific accounts of traditional Tuparí culture (1952; 1975). Caspar also collected extensive word lists of all the languages he encountered, including Djeoromitxí and Arikapú. He had access to Snethlage’s diaries and cited from them in his doctoral dissertation (Caspar 1953). The Czech linguist Loukotka had access to word lists by Snethlage and published a comparative account of them, grouping Arikapú and Djeoromitxí together with Mashubi as an isolated language family (Loukotka 1963). Fortunately, some of Snethlage’s and most of Caspar’s unpublished material from the Guaporé region is now becoming available online (Brijnen and Adelaar 2010).

5 According to the editor of this collection of letters, the manuscript map, which shows the geographic distribution of the Jê languages including the small western island of “Arikapú & Jabotí,” was included in a letter of 20 July 1934. It is likely, however, that it rather belonged to another letter, such as the next one from 22 January 1935, in which the Jê hypothesis was first mentioned. Furthermore, Snethlage did not leave the field before the end of November 1934 and probably showed his linguistic material to Nimuendajú first around the end of December in Belém.

6 As the result of an important digitalization initiative (2010) by Willem Adelaar and Hélène Brijnen of Leiden University, all of Caspar’s known unpublished linguistic data, as well as those of the nineteenth-century explorer Johann Natterer, and some of Snethlage’s and Sekelj’s, are being made available in the Digital Humanities section on the website of the Leiden University Library.
Traditionally, the indigenous peoples of the region must have had several thousand members each. Because of the absence of resistance against the infectious diseases of the Old World, more than half of the population of a group may die upon first contact with Westerners, if no protective action is taken. When Snethlage met the Arikapú, they were already a small group, divided over two settlements. In 1948, just before Caspar arrived in the region for the first time, the Yugoslavian journalist Tibor Sekelj traveled up the Rio Branco in order to visit the Tuparí. In his popular account, he mentions the neighboring groups as well and refers to the Arikapú as “the world’s smallest nation” (Sekelj 1950). In early 1954, the Indians of the Rio Branco suffered another horrible measles epidemic. The Arikapú were apparently hit very hard and when Caspar returned in 1954, they had dwindled to one small group. In 1968, the linguists Willem Bontkes (1968) and Robert Campbell surveyed the south of Rondônia for the evangelical Summer Institute of Linguistics. Bontkes recorded a short Arikapú word list at Rio Branco and mentioned the existence of 14 speakers. Campbell (1968) recorded a word list of Djeoromitxí there, from which some items were published in Rodrigues (1986).

In the 1980s, several anthropologists and linguists visited the region and met speakers of Jabutí languages. In her article, the ethno-historian Denise Maldi described southeastern Rondônia as a cultural area, called the Marico cultural complex, that includes Tupí groups, groups speaking isolated languages, and the Arikapú and Djeoromitxí (Maldi 1991). Her survey includes small word lists and traditional myths. Furthermore, the linguist Denny Moore of the Museu Goeldi collected comparative word lists in the Guaporé indigenous reserve, including Arikapú and Djeoromitxí (Moore 1988). Subsequently, the linguistics student Nádia Pires did fieldwork among the Djeoromitxí on the Guaporé reserve for her master’s thesis, which resulted in a concise description of Djeoromitxí (1992) and literacy material for the native language (Pires, Jabuti, and Jabuti 1994; 1995). The section on Jabutí in Dixon and Aikhenvald’s survey (1999:357–58) is based on the work by Pires. Between 2001 and 2004, Hein van der Voort conducted fieldwork with two known speakers of Arikapú and is preparing a description of the language (van der Voort [forthcoming b]). Certain lexical similarities between Jabutí languages and other languages of the region are mentioned in van der Voort (2005). In 2008, the linguistics student Thiago Vital conducted fieldwork in the Guaporé reserve and documented Djeoromitxí oral history and traditions on a series of DVDs and CDs. In the same year, linguistics students in Guajará-Mirim wrote dictionaries of Arikapú and Djeoromitxí, based on interviews with visiting native speakers, for their master’s theses (R. Ribeiro 2008 and M. Ribeiro 2008). In spite of these recent initiatives by linguists, we must emphasize that the documentation and analysis of both Arikapú and Djeoromitxí are still in an initial stage.
2.3. Characteristics of Jabutí languages. The Jabutí languages are genetically unrelated to the other languages and language families of the Guaporé region and present a number of unique features, both in their phonology and grammar (for comparative work on these languages, see van der Voort 2005; 2007). They are highly isolating languages, they have some phonological traits that are unusual in the region, and their lexicon is markedly different from that of the others. Nevertheless, they also share a number of characteristics with other languages, probably as the result of diffusion, especially with regard to the lexicon. Many loanwords are from Makuráp, which has functioned as a lingua franca for some time during the twentieth century (Caspar 1975:223, Sekelj 1950:50, and Snethlage 1937:127ff.).

When compared to each other, the Jabutí languages are highly similar. Since Arikapú and Djeoromitxí used to be closely neighboring languages, it is likely that they have diversified from a common ancestor in the same region. It seems less likely that they diversified somewhere outside of the region and at a certain moment immigrated together. This implies that the time depth between Arikapú and Djeoromitxí indicates the minimal length of time that the Jabutí languages exist in the Guaporé region. In the following subsections, we describe the basic characteristics of the Jabutí languages.

2.3.1 Arikapú. One important difficulty for the study of Arikapú is that the language has been out of use for a long time and the speakers who were recorded did not produce narrative texts or fluent dialogues. In principle, one cannot claim with absolute certainty that the syntactic patterns encountered in the recordings are original rather than the result of influence by more current languages, such as Djeoromitxí, Makuráp, Tuparí, or Portuguese. Also, certain morphological properties of the language may have become obsolete. Although the fluency of the remaining speakers was difficult to assess, one nevertheless gets the impression that the different speech samples they produced are consistent and highly similar in structure. In this respect, it is important to remember that the two remaining speakers lived far apart and did not maintain any contact with each other. Now, there is only one speaker of Arikapú left.

2.3.1.1. Phonology. According to recent analyses (van der Voort [forthcoming b]), the phonological system of Arikapú comprises eight oral vowels and six nasal vowels. The consonant system includes ten phonemes (see table 1).

Although Arikapú has a number of nasal vowels, which are confirmed by minimal pairs such as nìkra ‘hips’ vs. nìkrã ‘wrist’, takɔ̰ ‘it is hot’ vs. takɔ̣ ‘it is finished’, tari ‘to clear the field’ vs. tarĩ ‘it is dry’, and ku ‘to eat’ vs. kũ ‘to leave’, there is also some variation between nasal and oral pronunciation of vowels, and spread of nasality to neighboring (usually left) syllables. Hence there is limited evidence for a nasal phoneme /õ/, whereas most occurrences of the nasal phoneme /ẽ/ seem to be the result of nasal
spreading, except in two homophonous lexical roots, rē ‘egg’ and ‘dance’, which contrast minimally with re ‘cord, grub’. (For a likely diachronic explanation of this exception, see 4.1.1.)

The language also has diphthongs, some of which contrast with combinations of multiple vowels that belong to different syllables, as in kaj ‘head’ vs. ka.i ‘hair, fur, feathers’. Stress is placed on the ultimate syllable of the word. The /j/ occurs also in onset position, in words like mujo ‘long club’, kaju ‘wing, feather’, rajo ‘uncle’. Furthermore, there are some cases of /j/-insertion in stressed syllables in certain words, such as kamõ ‘hill’ vs. kamõtjutõ ‘on top of the hill’, kũ ‘grind’ vs. õhãkũ ‘grind seeds’, and inũũõ(j)nõ ‘I’m sleeping’ (see table 2).

Arikapú has some very frequent consonant combinations, such as /pr, kr, mr/, although their distribution is somewhat limited. Less common consonant combinations are /pw, kw, mw, pj, nj, wr, rn/.

There are some frequently occurring consonant alternations in Arikapú. The most common one concerns /m/, whose allophones [b] and [mb] occur in free variation with each other, before oral vowels, as in /miku/ [biku] ~ [mbiku] ‘yard’; /mo/ [bu] ~ [mbu] ‘arrow’. Word-internally, the preceding vowel is often nasalized, as in /kamu/ [kãbu] ~ [kam’bu] ‘young one’; /kra-
mrã/ [krãˈbɾã] ~ [kraˈmbrã] ‘beads’ (lit., ‘stone-little’) (['] indicates stress placement).

A similar pattern concerns the /n/, which has the freely alternating allophones [d] and [nd] before oral vowels and which, word-internally, causes nasalization of the preceding vowel, e.g., /nuni/ [ndundi] ~ [dundi] ~ [dũdi] ‘breast’.

Note that /m/ and /n/ also have nasal allophones, [m] and [n], respectively. These allophones occur mostly preceding nasal vowels and are in complementary distribution with [b/mb] and [d/nd]. Counterexamples to this complementarity usually involve loanwords. There are a few rare counterexamples in which [m] is attested preceding an oral vowel, but there always seems to be free variation between that oral vowel and its nasal counterpart, and no minimal pairs are attested.

The [dʒ] should probably not be considered as a separate phoneme, since its distribution is limited to preceding an [i] and does not contrast with [d] in that position. Although the [n] is also attested preceding an [i], that [i] appears to be in free variation with nasalized [i] in the same position. Therefore, all three should be regarded as allophones of /n/.

There is some phoneme fluctuation between /tʃ/ and /h/. These phonemes occur in free variation only in certain lexemes: tʃaroko ~ haroko ‘talk, say, language’. This alternation probably represents a relic of a similar, still productive alternation in Djeoromiti between /r/ and /h/ (see 4.2.1 below). There is variation between the presence and absence of /tʃ/ in a number of lexemes, such as i ~ tʃi ‘bone’ and pɔa ~ pɔtʃa ‘to form, shape, knit’.7

The glottal stop is probably not a phoneme but an automatic phenomenon that is motivated by a tendency to preserve the basic syllable structure CV.

Note that in the speech of one of our consultants, the /k/ is often aspirated [kh] when it precedes back rounded and central vowels. This is probably due to interference from Djeoromiti.

2.3.1.2. Grammar. From the available data it appears that the Jabutí languages are grammatically similar. They are relatively isolating languages. The basic parts of speech are nouns, verbs, adverbs, and particles. Adjectives are not distinguishable from uninflected intransitive verbs. Basic constituent orders are noun-modifier and SOV or OVS, which are often given as interchangeable optional alternatives by the consultants. There are a number of obligatory prefixes for person marking, the majority of which are distributed according to an ergative pattern. There is no case marking of core arguments, but there is a handful of postpositions that mark the semantic function of

7 In the case of i ~ tʃi ‘bone’, it is worth mentioning that the consonant-initial form occurs only as the second element in compounds, or preceded by a prefix. It is likely that, in such cases, the basic form is i, and that [tʃ] developed originally as transition glide.
oblique arguments. The Arikapú elements for person reference and their attested use are listed in table 3.8

The person prefixes can in principle be attached to members of any lexical category, be it a noun, verb, or adverb. When attached to nouns, they have a possessive sense, as in example (1) below, and when attached to verbs, they express a structural argument, as in (2).

(1) a-kanū
2-hat
‘your hat’

(2) ta-wērē
3-alone
‘He is by himself’.

When attached to intransitive verbs, the person prefixes indicate the subject. When attached to transitive verbs, the person prefixes indicate the object, as in (3) and (4). This suggests that Arikapú has an ergative person-marking strategy.

---

8The following abbreviations are used in this article: 1s = first-person subject; 1sg = first-person singular; 1pl = first-person plural; 2 = second person; 3 = third person; adj = adjectivizer; aik = Aikanã (isolate); aku = Akütsũ (Tuparí); api = Apinajé; ari = Arikapú; atr = advanced tongue root; ave = aversive; bor = Boróro; comit = comitative; ctfg = centrifugal direction; dat = dative; die = Djeoromix̆̄; do = direct object; fut = future; intens = intensifier; kal = Kaingang; kan = Kanoẽ (isolate); krah = Krahó; krj = Karajá; knk = Krenák; kwa = Kwaza (isolate); loc = locative; lp = linking prefix; mak = Makuráp (Tuparí); mek = Mekens (Tuparí); mxx = Maxakalé; n = noun; nom = nominalizer; ofy = Ofayé; paj = Proto-Amazonian Jê; p.c. = personal communication; pjab = Proto-Jabuti; pj = Proto-Jê; pl = plural; por = Portuguese; pos = possessive; prk = Parkatéjê; rel = relational (or linking) prefix; ri = Rikbaktsú; sal = Salamãi (Mondé); sg = singular; sp = species; suy = Suyá; trans = transitive marker; tup = Tuparí (Tuparí); v = verb; xav = Xavante; xer = Xerente; way = Wayurú (Tuparí).
(3) \textit{ahe=ne i-arā-tō}
\begin{align*}
\text{you}=2s & \text{1sg/3-see-NEG} \\
& \text{‘Don’t you even look at me!’}
\end{align*}

(4) \textit{tfi-koni-wiro}
\begin{align*}
\text{1pl-shoot-fut} & \\
& \text{‘(He) is going to shoot us’.}
\end{align*}

The Arikapu third-person intransitive argument prefix \textit{ta-}, which does not occur with transitive predicates, was also attested in possessive constructions. Both in form and function it is conspicuously similar to the coreferential/reflexive possessor (‘his own’) prefix in other Macro-Jê languages (Rodrigues 1986:55):

(5) \textit{ta-nīrō-nō ta-nūtō}
\begin{align*}
\text{3-place-loc} & \text{3-sleep} \\
& \text{‘He’s sleeping in his own house’.}
\end{align*}

Arikapú is a “PRO-drop” language, and the use of pronouns and other independent argument nouns is optional. Nevertheless, pronouns occur frequently in the data, usually in subject function. The following alternative phrases show the optionality of pronouns:

(6) \textit{ihe=nō i-pī}
\begin{align*}
\text{1s} & \text{1sg/3-dead} \\
& \text{‘I’m drunk’}.
\end{align*}

(7) \textit{i-pī=nō}
\begin{align*}
\text{1sg/3-dead}=1s & \\
& \text{‘I’m drunk’}.
\end{align*}

Apart from pronouns and prefixes, two enclitics are listed in table 3, =nō and \textit{=ne}. These enclitics agree respectively with the first- and the second-person subject of declarative, interrogative, and sometimes imperative sentences. As appears from (7), they may be used simultaneously with person prefixes, but they do not display ergative behavior and they seem to be optional. The most important distributional characteristics of subject clitics are that they tend to be attached to the first syntactic constituent of the sentence; they are never attached to object constituents; and they can be attached to oblique arguments. Note the alternative word orders in the following examples:

(8) \textit{ihe=nō tfi-rihe hē-wehe}
\begin{align*}
\text{1s} & \text{1pl-food cook-intens} \\
& \text{‘I’m making food’.}
\end{align*}
The position of the subject clitic in (9) implies furthermore that the combination of object and predicate forms a syntactic constituent.

2.3.1.3. Nonspecific use of personal prefixes. The prefixes for first-person singular and plural, \( i- \) and \( tf\)-, are homophonous with the nonspecific or impersonal prefixes. Because of its ambiguity and its not entirely resolved status, \( i- \) is glossed as ‘1sg/3’. Because impersonal use of first-person plural is not an uncommon phenomenon across languages, we will continue to gloss \( tf\)- as ‘1pl’, assuming that it represents the same morpheme. The two prefixes can occur as third-person prefixes that in possessive constructions seem to function often as a dummy possessor and with transitive verbs or semantically adjectival intransitive verbs as a dummy argument. There is a distributional difference between them and the normal person prefixes, which can occur in combination with an overt pronoun. When the possessor or the third-person object is expressed by a full noun, the impersonal prefixes are absent.\(^9\) The following examples contrast possessive constructions:

(10) \( tfi\)-praj
    1pl-foot
    ‘our foot’

(11) \( n\o wa\)-praj \( ar\o =n\o \) (ihe)
    tapir-foot see=1s I
    ‘I saw a tapir track’.

(12) \( i\)-tfawa
    1sg/3-flower
    ‘(its) flower’, ‘my flower’

(13) \( kuk\o \)-tfawa
    jambu-flower
    ‘jambu flower’

(12) shows that the absence of a (pro)noun may allow for grammatical ambiguity, since, in principle, various interpretations are possible. The impersonal prefix \( tf\)- is often interpreted as a plural third person with transitive verbs:

\(^9\) It deserves mention here that there are no third-person pronouns in Arikapú.
On some occasions, there was apparently no pragmatic difference between the prefixes *i-* and *tSi-*, whether in subject function or in object function:

(15) \( \text{ah} \text{e}=\text{n}\text{e} \quad \text{i-kuraj} \)
\[ \text{you}=2\text{s} \quad 1\text{SG}/3\text{-throw} \]
‘Did you already throw (it) away?’

(16) \( \text{ah} \text{e}=\text{n}\text{e} \quad \text{tSi-kuraj} \)
\[ \text{you}=2\text{s} \quad 1\text{PL}\text{-throw} \]
‘Did you already throw (it) away?’

The prefix *tSi-* strongly resembles another prefix, *tfe-*. It does not occur on intransitive verbs, but it was attested in a possessive function once, and it functions mostly as an unspecified direct object prefix for a transitive verb.

(17) \( \text{ah} \text{e}=\text{n}\text{e} \quad \text{pat}\text{f}\text{i} \quad \text{t}=\text{n}\text{e}, \quad \text{a-naj} \quad \text{tfe-iro} \)
\[ \text{you}=2\text{s} \quad \text{tobacco} \quad \text{bring}=2\text{s} \quad 2\text{-older.brother} \quad \text{do-want} \]
‘Did you bring tobacco? Your older brother wants some’.

The prefix *i-* with an impersonal sense resembles the dummy element *i-*, or *e-* in a number of Guaporé languages of different affiliation (see van der Voort 2005). However, it remains a question whether the impersonal prefixes *i-* and *tSi-* really should be considered as different prefixes that happen to be homophonous with the first-person singular and plural prefixes *i-* and *tSi-* . It seems equally likely that they represent impersonal use of the category first person, similar to impersonal use of the second-person singular or the first-person plural in many other languages.

2.3.1.4. Postpositions. Arikapú has a limited set of elements that indicate the specific semantic function of the oblique arguments they follow. These elements indicate roles such as aversive, dative, instrumental, locative, simulative, comitative, and ablative.

(18) \( \text{ih} \text{e}=\text{n}\text{ô} \quad \text{i-kon}\text{ô} \quad 3\text{waw} \quad \text{rij} \)
\[ \text{I}=1\text{s} \quad 1\text{SG}/3\text{-angry} \quad \text{João} \quad \text{AVE} \]
‘I am angry with João’.

(19) \( \text{tfuk}\text{ô} \quad \text{ü} \quad \text{i-ri} \)
\[ \text{rope give} \quad 1\text{SG}/3\text{-DAT} \]
‘Give me rope!’
2.3.1.5. Other grammatical elements. Arikapú has grammatical elements which we have considered as derivational suffixes for the time being, such as nominalizer/adjectivizer -hã, diminutive -mraj, augmentative -tfí and -rukre, intensifier -wehe, negative imperative -pi, future -wiro, and purposive -wə. Note the following example:

(20) tfí-hã kari-hã iro=nô ihe
maize-kernel dry-nom want=1s 1
‘I want toasted maize’.

In addition, there are elements that resemble nominal classifiers, such as -hã ‘seed, kernel’, -ka ‘round’, -ko ‘skin, shell’, -ko ‘hole’, -mrê ‘porridge, dregs’, -nê ‘flesh’, -nê ‘thorn, leaf, cord’, -nũ ‘porridge, paste’, re ‘wormlike’, -rê ‘egg’, and -mrê ‘powder, flour’. These elements were rarely attested by themselves and were usually preceded by nouns or person prefixes, as in:

(21) mure-mrê
manioc-porridge
‘manioc porridge’

(22) a-mrê
2-porridge
‘your porridge’

For the expression of their semantic content only, the prefix i- is usually applied, as in imrê ‘porridge’, ihã ‘seed’, iko ‘hole’, etc. In these respects, the classifier-like elements strongly resemble those attested elsewhere in the Guaporé region and could be attributed to areal diffusion. Some of the classifying forms even occur in almost identical form in unrelated languages, such as Kwaza (isolate) -mê ‘porridge’, -nê ‘thorn’, -nũ ‘porridge, paste’ (van der Voort 2005:397). Since the majority of the classifier-like elements of Arikapú are reconstructible in Proto-Jabutí and Proto-Jê, as will be shown in 2.4.3 and 4.1, their origin is most likely Macro-Jê.

2.3.2. Djeoromitxí. As mentioned in 2.2, Djeoromitxí is a living language that still is transferred to the younger generations. Its speakers are very much in favor of keeping the language alive and actively support its further development in terms of expanding its range of use beyond the traditional context. The literacy project set up by Nádia Pires has been quite successful and Djeoromitxí is used sometimes as a vehicle for teaching in

---

10 One of the classifiers that is highly productive in the isolated languages of the Corumbiara and Pimenta Bueno region, Aikanã, Kanoẽ, and Kwaza, the classifier -mũ ‘liquid’ is encountered in several river names in use by the Arikapú and Djeoromitxí; that probably means that these hydronyms were borrowed by the Jabutí languages from one (or several) of the other languages.
primary schools on the reserve. Even modern-style, romantic, guitar-accompanied songs are created and performed in the language. As long as the speakers are allowed to continue as a community, the language will probably survive.

2.3.2.1. Phonology. According to the analyses by Pires (1992), the phonological system of Djeoromitxí comprises seven oral vowels, four nasal vowels, and 14 consonants. In this article, however, we do not consider all of those units as phonemic, and count only 12 consonants (see table 4). The close central rounded vowel /Á/ is often pronounced as [ø].

There are no diphthongs in Djeoromitxí and combinations of multiple vowels always belong to different syllables. The glottal stop [ʔ] is an optional phonetic phenomenon at the encounter of two vowel phonemes. Stress is placed on the ultimate syllable of the word (see table 5).

Although the available data on Djeoromitxí are less decisive, we assume that the language is essentially like Arikapú in that it does not have a set of

---

Although Pires’s data and analyses are reliable and systematic, they are not complete, and additional data obtained in the field have in certain cases presented us with a different perspective. That is not to say that we consider the analysis put forward here as definitive. Djeoromitxí merits a much more wide-ranging and profound study than has been done so far, especially since there is still a viable speakers’ community to make that possible.

---

### TABLE 4
**Djeoromitxí Vowel Phonemes**

<table>
<thead>
<tr>
<th></th>
<th>Front</th>
<th>Central</th>
<th>Back</th>
<th>Front</th>
<th>Central</th>
<th>Back</th>
</tr>
</thead>
<tbody>
<tr>
<td>Close</td>
<td>i</td>
<td>u</td>
<td>u</td>
<td>i</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Close-mid</td>
<td>o</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Open-mid</td>
<td>e</td>
<td>o[ɔ]</td>
<td>e</td>
<td>o[ɔ]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Open</td>
<td>a [a]</td>
<td></td>
<td>a</td>
<td>a</td>
<td>[a]</td>
<td></td>
</tr>
</tbody>
</table>

### TABLE 5
**Djeoromitxí Consonant Phonemes**

<table>
<thead>
<tr>
<th></th>
<th>Labial</th>
<th>Alveolar</th>
<th>Velar</th>
<th>Glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>− Voice</td>
<td>+ Voice</td>
<td>Voice</td>
<td>Voice</td>
</tr>
<tr>
<td>Plosive</td>
<td>p</td>
<td>t</td>
<td>k</td>
<td></td>
</tr>
<tr>
<td>Nasal</td>
<td>m</td>
<td>n</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trill/tap</td>
<td>r</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fricative</td>
<td>h</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Affricate</td>
<td>ps</td>
<td>bz</td>
<td>tf</td>
<td>dʒ</td>
</tr>
<tr>
<td>Approximant</td>
<td>w</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
voiced plosive consonant phonemes that are distinct from nasal consonants. The distribution of [b] and [d] versus [m] and [n] appears to be largely complementary, [b] and [d] occurring basically only before oral vowels, and [m] and [n] before either nasal or oral vowels.12

The affricate /bz/ occurs only in one specific environment, before a close front vowel /i/ (see n. 16 for a historical explanation). It nevertheless contrasts with /m/. The vowel that follows /bz/ is often not pronounced, as in /bzirE/ [bzirc] ~ [bzrc] ‘capuchin monkey’. Since no minimal pair was encountered, the phonemic status of /ps/ is less obvious, although it contrasts with /p/ in the same environment.

The /dʒ/ is a phoneme that does not occur before /a/ and /ə/. In the practical orthography created by the Djeoromítxí and Nádia Pires, it is written as a digraph <dj>, similar to the /tʃ/, which is written as <tx> (Pires, Jabuti, and Jabuti 1994; 1995).

Note that preceding the close rounded central vowel [u], the /k/ is pronounced as [q]. Preceding the central vowel /ə/ and the back vowel /u/, it is often aspirated [kh]. Preceding back vowels and [u], the /p/ is usually pronounced as a bilabial fricative [ɸ], or as an affricate [pɸ].

There is a morphophonological alternation that is characteristic of Djeoromítxí. When /h/-initial words are preceded by a personal prefix, or when they occur as the second element in a composition, the /h/ is replaced by /t/. This is shown in the following examples:13

(23) huu humti
     I   be.ill
     ‘I’m ill’.

(24) huu pa=rumti
     I  foot=be.ill
     ‘My foot hurts’.

(25) hõmeka=rumti
     head=be.ill
     ‘I have a headache’.

12 Partially similar systems were attested in languages of the Tuparí language family, like Makuráp and Mekens (Moore and Galucio 1994), and in a number of Jê languages (Rodrigues 1999). The minimal pair registered in Pires (1992:37) for the opposition between [d] and [n] is not convincing since it contains an unanalyzed phrase <noko> ‘to eat’ (/nõ ku/, lit., ‘eat food’) and a loanword from Makuráp: <doko> ‘electric eel’ (mak [‘nduku], Die miraTiti). More research is necessary, however.

13 Examples such as (24) and (25) are highly productive. Although the theoretical significance of these constructions has not yet been fully explored, they constitute an interesting grammatical parallel with Proto-Jê, where similar noun–verb compounds can be reconstructed.
the jabutí language family

(26) hamə
‘He is tired’.

(27) hɯ hamə
I’m tired
‘I’m tired’.

(28) adzə a-ramə
you 2-tired
‘You’re tired’.

Pires (1992:46) contains data that indicate that there is similar morphophonological alternation between /h/ and /n/ with certain roots. These alternations remind one of the so-called relational prefixes described for some Macro-Jê and Tupí languages (e.g., Rodrigues 1999 and Seki 2000), as discussed in 3.3 and 4.3. If, as the phonological correspondences seem to suggest (see 4.3), such alternations in Djeoromitxí are indeed related to the ones found in Jê and other Macro-Jê families (Karajá and Ofayé, for instance), this would constitute an important piece of evidence for genetic relationship (a “shared aberrancy”).

2.3.2.2. Grammar. The above examples suggest a high degree of similarity between the grammatical structures of Djeoromitxí and Arikapú. As mentioned in the previous section, the basic parts of speech of the Jabutí languages are nouns, verbs, adverbs, and particles, while adjectives14 are descriptive intransitive verbs. Basic constituent orders are noun-modifier and SOV or OVS. Djeoromitxí also has obligatory prefixes for person marking, albeit that the forms involved differ from Arikapú. Compare table 3, which shows Arikapú person reference, to table 6, which shows the Djeoromitxí forms and their attested use.

14 As identified by Pires (1992).
In Djeoromitxí, first-person singular is unmarked, whereas the first-person plural and the second-person have obligatory prefixes. The third-person is marked by a prefix when an overt argument or possessor noun is absent. The use of personal pronouns is in principle optional. The following examples show possessive use of the person prefixes:

(29) **nik**
    field
    ‘my planted field (with maize)’

(30) **i-rawa**
    3-flower
    ‘(its) flower’

Djeoromitxí also has an alternative way to express possession, through the possessive particle *dzeweta*. This particle can receive person prefixes. Without any prefixes it means ‘mine’ or ‘my’:

(31) **dzeweta**    **tiru**
    pos    husband
    ‘my husband’

(32) **a-dzeweta**    **kau**
    2-pos    neck.hair
    ‘your hair in the back of the neck’

In constructions like these, as well as in other possessive constructions with a possessor noun, the possessed head does not receive person prefixes.

Djeoromitxí has an ergative person-marking strategy. When attached to intransitive verbs, person prefixes indicate the subject; when attached to transitive verbs, person prefixes indicate the object. In the following examples, person marking, or its absence, indicates the subject of the intransitive predicate:

(33) **hu**    **hakmente**
    I    yawn
    ‘I’m yawning’.

(34) **hiru**    **hi-hörahi**
    we    1PL-drunk
    ‘We’re all drunk’.

With transitive predicates, the object rather than the subject is indicated by person prefixes, the distribution of person marking being ergative in Djeoromitxí. The third person object is marked by the prefix **i-**.
(35) *hu a-təmi adʒe*
   I 2-beat you
   ‘I’m going to beat you’.

(36) *hu i-te a-ri i-ū adʒe*
   I 3-bring 2-DAT 3-give you
   ‘I brought it to give to you’. 15

Note that Djeoromitxí is not considered as an ergative language in Pires (1992). The behavior of the prefix *i*- is explained as a result of complex word formation. A more recent version of this view is based on the analysis of transitive objects as subjects in passive constructions (Pires 1998). Although this nominative–accusative analysis of Djeoromitxí may also be plausible, it merits further investigation and is not adhered to here.

2.3.2.3. Nonspecific use of person prefixes. Unlike the other person prefixes, the *i*- prefix does not appear when the argument is represented by an overt noun or pronoun. Like in Arikapú, this prefix often represents a nonspecific or indefinite third-person argument. The difference with Arikapú is that *i*- is never interpreted as a first person in Djeoromitxí. Note the following examples:

(37) *bziru kuhi*
    water spill
    ‘spill water’

(38) *i-kuhi-tõ*
    3-spill-NEG
    ‘Don’t spill it!’

Like the Arikapú prefix *tʃi*- , the Djeoromitxí prefix *hi*- can be interpreted both as a first-person plural prefix and as an impersonal prefix. This is probably a cross-linguistically common strategy. Not surprisingly, it is often encountered with human body-part nouns:

(39) *hi-məhiri*
    1PL-lung
    ‘one’s lungs’

(40) *hi-opədzį*
    1PL-adam’s.apple
    ‘one’s Adam’s apple’

15 Constituent order *ari iū adʒe hu ite* was also attested.
The metaphorical extension of the first-person plural to impersonal is clearest in examples that cannot be so easily interpreted literally, such as (40), which was uttered by a female consultant.

2.3.2.4. Postpositions. Oblique arguments are marked for different roles, such as locative, instrumental, comitative, by a closed set of postpositions:

(41) $i$-hi $pakə$ $i$-pu

3-hair COMIT 3-eat

‘(He) eat(s) game with hair and everything’.

(42) nipɛru $tʃɛ$ a-run$ə$

bench LOC 2-sit

‘You sit on the bench’.

Pires (1992) has furthermore documented postpositions such as dative $ma$, agentive $me$, and beneficiary $eni$.

2.3.2.5. Other grammatical elements. Even though the Jabutí languages are relatively isolating, there is some morphology apart from bound person marking. Djeoromitxí has derivational suffixes such as nominalizer -tʃi, diminutive -titi (Pires 1992), augmentative -tʃitʃi, and intensifier -wehɛ.

Furthermore, there are grammatical elements of which it is hard to say whether they are free particles or bound morphemes, such as negative $tɔ$, interrogative $hi$, future $ma$ (probably, the same as the dative postposition), past $tʃe$, and nominalizers $a$, $hɔ$, and $ɛ$ in Pires (1992). (43) is an example with the adjective nominalizer $hɔ$:

(43) $tʃitʃi$-h$ɔ$ $kurɨ$-h$ɔ$ $u$

maize-kernel dry-ADJ give

‘Give me toasted maize’.

Finally, there are nominal classifiers, such as $-dɛ$ ‘egg’, $-hɔ$ ‘seed, kernel’, $-ká$ ‘round’, $-kɔ$ ‘skin, shell’, $-kʉ$ ‘hole’, $-má$ ‘porridge’, $-nĩ$ ‘flesh’, $-nĩ$ ‘thorn, leaf, cord’, $-rɔ$ ‘small’, $-rɛ$ ‘wormlike’, and $-tu$ ‘powder, flour’. Some of these are analyzed by Pires (1992) as the second part of a nominal compound. They were not attested as independent elements. They clearly have derived from nouns, and some of them are still attested as nouns, most of which are obligatorily possessed (with the exception of $kɔ$ ‘skin, shell’). If their semantic content is to be expressed as an independent form at all, these elements minimally require the prefixation of $i$-, as in $ihɔ$ ‘kernel’, $i$k$ʉ$ ‘hole’. The classifier-like elements of Arikapú and Djeoromitxí constitute highly similar sets of sometimes rather unproductive morphemes that belong to a
small closed category with a quite limited distribution. To a certain extent their behavior and form resemble those of a number of unrelated languages of the region, which could be explained by areal diffusion. Most of them, however, can be reconstructed for Proto-Jabutí (see 2.4.3) and have lexical cognates in other Macro-Jê families (see 4.1), which suggests that, if areal diffusion of classifiers took place, (Proto-)Jabutí may have been a major source.

2.4. Proto-Jabutí. There is little doubt that Arikapú and Djeoromitxí are genealogically closely related. The similarities on the lexical, phonological, and grammatical levels are very strong and a great number of words can be reconstructed. In this section, we present evidence for the relationship by reconstructing the phonological inventory and a number of lexical and grammatical forms of the proto-language.

2.4.1. The phonological system. The lexical database on which this section is based contains approximately 1,410 unique Arikapú roots and 1,095 unique Djeoromitxí roots. Among these, around 1,085 sets of entries were comparable. After sets that showed no formal similarity whatsoever were excluded, and after probable loans and onomatopoeic forms were separated, roughly 500 sets of entries remained that are possibly cognate. On the basis of the regular correspondences between the phonemes in similar positions in the forms, approximately 200 proto-forms were fully reconstructed, as well as the shared ancestral phonological system. Table 7 shows the reconstructed vowel phonemes of Proto-Jabutí.

In the following lists, regular correspondences between Arikapú and Djeoromitxí phonemes are shown, with their reconstruction in Proto-Jabutí. The numbers correspond to relevant items in the appended Swadesh basic 100-word list. Apart from those items, there are many other fully corresponding cognate sets to support the postulated regular correspondences. The complete list of cognate sets and full information on the conditioning
environments of the reconstructed phonemes are given in van der Voort (2007).

Table 8 shows the reconstructed consonant phonemes of Proto-Jabutí, and below that, lists of regular correspondences and references to items in Appendix A.
Under Pires’s (1992:20) analysis, /bz/ represents a rare phoneme in Djeoromitxí that occurs only before /i/, where it contrasts with /m/: /bz(i)ru/ ‘water’ vs. /mi/ ‘genipap’. In Arikapú cognates, this contrast does not exist, but there is a contrast between /i/ and /i/ in the same position: /mi/ ‘water’ vs. /mi/ ‘genipap’, respectively. Probably, in an earlier phase of Djeoromitxí, /bz/ was an allophone of /m/ conditioned by a high “tense” vowel /i/; whereas the allophone /m/ occurred before the high “lax” vowel /i/. When the contrast between these vowels was lost in Djeoromitxí, /bz/ must have acquired phonemic status. There are only a few complete cognates involving /bz/, e.g., Proto-Jabutí *mihí, ari miįfį, die bzihi pikorų ‘monkey species’ and Proto-Jabutí *tʃamųj, ari tʃamųj, die habzi ‘cotton’.

In some cases, such as *ni ‘to salt’, *n disappeared in Djeoromitxí, changing the word to i, whereas it was preserved in Arikapú: ni.

---

### TABLE 8
PROTO-JABUTÍ CONSONANT PHONEMES

<table>
<thead>
<tr>
<th></th>
<th>Labial</th>
<th>Alveolar</th>
<th>Velar</th>
<th>Glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plosive</td>
<td>p</td>
<td>t</td>
<td>k</td>
<td></td>
</tr>
<tr>
<td>Nasal</td>
<td>m</td>
<td>n</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trill/tap</td>
<td>r</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Voiceless affricate</td>
<td>tʃ</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Voiced affricate</td>
<td>dʒ</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Approximant</td>
<td>w</td>
<td>j</td>
<td>h</td>
<td></td>
</tr>
</tbody>
</table>

---

16 Under Pires’s (1992:20) analysis, /bz/ represents a rare phoneme in Djeoromitxí that occurs only before /i/, where it contrasts with /m/: /bz(i)ru/ ‘water’ vs. /mi/ ‘genipap’. In Arikapú cognates, this contrast does not exist, but there is a contrast between /i/ and /i/ in the same position: /mi/ ‘water’ vs. /mi/ ‘genipap’, respectively. Probably, in an earlier phase of Djeoromitxí, /bz/ was an allophone of /m/ conditioned by a high “tense” vowel /i/, whereas the allophone /m/ occurred before the high “lax” vowel /i/. When the contrast between these vowels was lost in Djeoromitxí, /bz/ must have acquired phonemic status. There are only a few complete cognates involving /bz/, e.g., Proto-Jabutí *mihí, ari miįfį, die bzihi pikorų ‘monkey species’ and Proto-Jabutí *tʃamųj, ari tʃamųj, die habzi ‘cotton’.

17 In some cases, such as *ni ‘to salt’, *n disappeared in Djeoromitxí, changing the word to i, whereas it was preserved in Arikapú: ni.
The following items suggest that Proto-Jabutí also has glides. The *j usually functions as a semivowel in falling diphthongs. Pires (1992:42–44) notes that there are no diphthongs on the phonological level in Djeoromitxí. Some examples indicate that the *j was preserved only in Arikapú. Other examples suggest optional /j/-insertion in Arikapú:

<table>
<thead>
<tr>
<th>PROTO-JABUTÍ</th>
<th>ARIKAPÚ</th>
<th>DJEOROMITXÍ</th>
<th>SWADESH LIST NUMBER</th>
</tr>
</thead>
<tbody>
<tr>
<td>*j</td>
<td>j</td>
<td>ɐ</td>
<td>8, 45, 46, (65), 66, 67</td>
</tr>
<tr>
<td>*j</td>
<td>j</td>
<td>i</td>
<td>(see van der Voort 2007)</td>
</tr>
<tr>
<td>*w</td>
<td>w</td>
<td>w</td>
<td>85</td>
</tr>
<tr>
<td>*w</td>
<td>w</td>
<td>u, ɐ</td>
<td>/a, ɐ_</td>
</tr>
</tbody>
</table>

As shown in 2.3.1.1, Arikapú has a number of lexemes in which /tʃ/ idiosyncratically alternates with /h/ and with absence of /tʃ/. We showed in 2.3.2.1 how certain lexical roots in Djeoromitxí are subject to morphophonological alternation between /tʃ/ in onset position word-internally and /h/ word-initially. These two alternating sets of consonants may be related historically but not synchronically. Be they verbs or nouns, many Djeoromitxí roots that display this morphophonological alternation, such as -rua, hua ‘to sing’, are cognate with roots in Arikapú that have /tʃ/ in any corresponding position (i.e., both with or without a preceding prefix or noun), in this case tʃua, reconstructible for Proto-Jabutí as *tʃua.

Many Arikapú roots with /tʃ/, some of which display alternation with [ʃ], like pʃa, pʃʃa ‘to form, put, weave’, are cognate with roots in Djeoromitxí that have nonalternating /r/ (word-internally) or /h/ (word-initially) in a corresponding position, in this case pera, reconstructible in Proto-Jabutí as *pʃʃa.

Before /i/, as in Proto-Jabutí *pudʒi ‘needle’, *dʒ corresponds to /dʒ/ in Djeoromitxí, in this case pudʒi, Arikapú puti.

Proto-Jabutí *r underwent a split in Djeoromitxí: before oral vowels it remained as /r/ and before nasal vowels it changed to /n/ and merged with Proto-Jabutí *n, as in *rʃi ‘braise, grill’, Dje rʃi, Ari rʃi.

Arikapú has a number of consonant combinations with /r/ in second position in the onset, as in kumʃʃ ‘peanut’, which are absent in Djeoromitxí, as in kumʃ ‘peanut’. Assuming that these consonant combinations underwent simplification in Djeoromitxí, they can be reconstructed for Proto-Jabutí, in this case *kumrʃ.

In a few cases, [ʃʃ] was preserved both in Arikapú and in Djeoromitxí, although the cognates correspond only partially or are from possible loanwords: *hāiʃ, ARI hātʃ, Dje hōtʃi ‘slow’; and *meʃʃa, ARI mejʃʃa, Dje melʃʃia ‘tayra’.

See 2.3.1.1. In such cases, reconstructed *j is placed in between brackets.

Few examples exist of *w used as a semivowel, as in *kaw ‘hair in the back of the neck’. In this case, Djeoromitxí was simplified or made use of an allophone of /ul/, as in kau, Arikapú kaw.
A considerable number of lexical and grammatical roots of the Jabutí languages are cognate with roots of Jê languages. These are discussed in Appendix A.

2.4.2. The lexicon. Some parts of the lexicon are more suitable than others for the purpose of historical comparison. In the above sections, the highest value is attached to items that represent the most basic concepts that can be expected to be lexicalized in each language, such as body parts, close kinship relations, and geographic phenomena, which can be found in Swadesh’s 100 basic word list. It is likely that such items are inherited from previous stages of a language, since they do not tend to represent unknown concepts for which new terms have to be introduced by borrowing or invention.

Although symbolic forms and borrowed forms may show similarities that confirm the same historical correspondences as basic forms do, they are also able to distort the picture and display similarities that are otherwise not accounted for in the specific reconstruction at hand. In the case of the Jabutí languages, there is no clear correspondence between Arikapú /tʃ/ and Djeoromixi /t/, for example. The only evidence for this correspondence is a set of items that may be sound-symbolic in origin: Arikapú katʃi vs. Djeoromixi kati ‘itch’. In van der Voort (2005:384), a number of possible sound-symbolic forms are listed that are similar to forms in neighboring languages, such as Arikapú kutaj, kokoro vs. Djeoromixi kutʃi, krukru ‘cicada’, which is similar to Aikanã and Kwaza (isolates) kuku, and Arikapú pupu vs. Djeoromixi popo ‘owl’, which is similar to Aikanã pupur ‘owl species’, Kwaza (isolate) s̄uβuβúd̄i- ‘(owl is) calling’ and Mekens (Tuparí) popoba ‘owl’. Due to their possible onomatopoetic origin, these forms cannot function as evidence for a genetic relationship.

An example that should not be used to confirm reconstruction of *w as a possible semivowel is Arikapú tfaw vs. Djeoromixi tfaw ‘stingray’. These forms are conspicuously similar to Makuráp tfaw, Mekens tʃaraw, Kanoẽ tsakɔwɔ, and Kwaza tsakarɔ ‘stingray’. In fact, this item may be one of a specific group of lexical items that are found throughout the region. Those items were discussed in van der Voort (2005:385–88) as possible indicators of a linguistic area. The following are also encountered in the Jabutí languages:26

ARI and DJE tʃiʃi ‘maize’: AKU atiti; KAN atiti; KWA atʃiʃi; MAK atiti; MEK atʃiʃi; TUP atiti; WAY atiti;
ARI pɔwru and DJE mioro ‘woodpecker’: KWA mauru; WAY mɔ̄ru
ARI tʃu and DJE nu ‘marico bag’: AK dyi; KWA súi; MAK etʃi

26 Most forms are from the fieldwork of van der Voort, except for the Mekens forms, which are from Galucio (2001), and some Kanoẽ forms, which are from Bacelar (2004).
ARI $tʃurim$ and DJE $hemi$ ‘potato’; 27 AÍK $tfidmu$ ‘pumpkin’, KWA $ṣurim$ ‘potato’, and SAL $ṣerim$ ‘pumpkin’
ARI $warwar$ and DJE $wirwir$ ‘star’: Aymara warawara (Willem Ade-laar, personal communication); KAN $wariwari$; KWA $waruwaru$; MEK $waruwaru$; TUP and MAK $waruwaru$

Obviously, none of these forms should be reconstructed to serve as evidence for a genealogical relationship between Arikapú and Djeoromitxí. Then again, even if certain items originate from borrowing or onomatopoeia, they may have been borrowed or created in Proto-Jabutí, before the split between the two Jabutí languages. The value of the reconstruction of such items lies in their confirmation of already independently identified and attested patterns of correspondence.

2.4.3. Grammar. Since it follows from 2.3.1 and 2.3.2 that the grammars of the Jabutí languages are rather similar, it seems reasonable to assume that their shared grammatical traits reflect aspects of the grammar of Proto-Jabutí. In the first place, the person-marking systems are similar. Both languages employ similar person prefixes to express subject or object, according to an ergative alignment pattern. In Proto-Jabutí the forms of two markers may have merged, creating homophony between the first and third person. This ambiguity was preserved in Arikapú, whereas in Djeoromitxí a zero-form emerged representing the first person:

<table>
<thead>
<tr>
<th>PERSON</th>
<th>PROTO-JABUTÍ</th>
<th>ARIKAPÚ</th>
<th>DJEOROMITXÍ</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>*i</td>
<td>i</td>
<td>~</td>
</tr>
<tr>
<td>2</td>
<td>*a</td>
<td>a</td>
<td>~</td>
</tr>
<tr>
<td>3</td>
<td>*i</td>
<td>i, Ø</td>
<td>~</td>
</tr>
<tr>
<td>1PL</td>
<td>*hi</td>
<td>tfi</td>
<td>~</td>
</tr>
</tbody>
</table>

Also, the nonspecific function of the first-person plural presented in sections 2.3.1.3 and 2.3.2.3 above is a shared trait, although there is also room for a universalist explanation.

Another shared trait is the use of similar, reconstructible, postpositional elements with similar semantic functions, as shown in 2.3.1.4 and 2.3.2.4. Not all attested postpositions were attested in both languages, however, such as Djeoromitxí agentive $me$ and dative $ma$, and Arikapú aversive $ri$ and simulative $r$:

27 The claim by van der Voort (2005:388) that the ARI and DJE forms would be cognate is not substantiated by the comparative evidence put forward in this article.
Other grammatical elements were discussed in 2.3.1.5 and 2.3.2.5. Some can be reconstructed for the proto-language:

<table>
<thead>
<tr>
<th>Function</th>
<th>Proto-Jabutí</th>
<th>Arikapú</th>
<th>Djeoromitxí</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ablative</td>
<td>*kunĩ</td>
<td>-kunĩ</td>
<td>kunĩ</td>
</tr>
<tr>
<td>Beneficiary</td>
<td>(*enĩ)</td>
<td>-nẽ</td>
<td>eni</td>
</tr>
<tr>
<td>Comitative</td>
<td>*pakəj</td>
<td>-pakəj</td>
<td>pakə</td>
</tr>
<tr>
<td>Dative</td>
<td>*ri</td>
<td>-ri</td>
<td>ri</td>
</tr>
<tr>
<td>Instrumental</td>
<td>*nɔ</td>
<td>-nɔ</td>
<td>ɔ</td>
</tr>
<tr>
<td>Locative</td>
<td>*tfɛ</td>
<td>-tfɛ</td>
<td>tfɛ</td>
</tr>
</tbody>
</table>

Negative        | *tɔ          | ɔ̅̊       | ɔ̅̊          |
Intensifier     | *-wehe       | -wehe    | -wehe       |
Augmentative    | *-tfifiant   | -tfifiant| -tfifiant    |

However, a large number of grammatical morphemes cannot be reconstructed. This may involve morphemes that were attested in one of the languages only, such as the Arikapú negative imperative -pi, or the Djeoromitxí interrogative hi and past tfɛ. Furthermore, some morphemes are very different from one language to the other, such as Arikapú future -wiro and Djeoromitxí future ma (a likely result of the grammaticalization of the dative postposition), Arikapú diminutive -mrəj and Djeoromitxí diminutive -tə or titi, and Arikapú purposive -wə(j) and Djeoromitxí purposive dzɛ. Finally, the languages may have influenced one another or may have been influenced by neighboring languages. A case in point is the perfective and adjective nominalizer, Arikapú hã (pronounced as [hã] or [hã]) and the Djeoromitxí adjective nominalizer ha, the reconstruction of which is impeded by the absence of systematic vowel correspondence. In Arikapú, however, the morpheme does not seem to be very productive, and its application in certain cases resembles a calque on Djeoromitxí.²⁸

Remember that the description of Djeoromitxí is far from complete and that Arikapú is a moribund language. It may be that future research in the field will produce more complete and reliable data, but it is also possible that the languages have lost certain categories over time or due to obsolescence.

As shown in 2.3.1.5 and 2.3.2.5, the Jabutí languages do seem to employ a limited set of classifying elements. Some of them can be reconstructed:

²⁸Moreover, its form is conspicuously similar to the nominalizer -hã of Kwaza (isolate), which may or may not be a coincidence.
<table>
<thead>
<tr>
<th>Meaning</th>
<th>Proto-Jabutí</th>
<th>Arikapú</th>
<th>Djeoromitxí</th>
</tr>
</thead>
<tbody>
<tr>
<td>seed, kernel</td>
<td>*-hã</td>
<td>-hã</td>
<td>-hõ</td>
</tr>
<tr>
<td>round</td>
<td>*-ka</td>
<td>-ka</td>
<td>-ka</td>
</tr>
<tr>
<td>skin, bark</td>
<td>*kʰ</td>
<td>kʰ</td>
<td>kʰ</td>
</tr>
<tr>
<td>shell</td>
<td>*-kakʰ</td>
<td>-kakʰ</td>
<td>-kakʰ</td>
</tr>
<tr>
<td>hole</td>
<td>*-ko</td>
<td>-ko</td>
<td>-kʰ</td>
</tr>
<tr>
<td>flesh</td>
<td>*-nɨ</td>
<td>-nɨ</td>
<td>-nɨ</td>
</tr>
<tr>
<td>thorn</td>
<td>*-nɨ</td>
<td>-nɨ</td>
<td>-nɨ</td>
</tr>
<tr>
<td>food, porridge</td>
<td>*nũ</td>
<td>nũ</td>
<td>nũ</td>
</tr>
<tr>
<td>wormlike</td>
<td>*-rɛ</td>
<td>-rɛ</td>
<td>-dʒɛ</td>
</tr>
</tbody>
</table>

Some classifiers resist reconstruction, such as Arikapú *-mrɛ, Djeoromitxí *-má ‘porridge, dregs’; Arikapú *-mrɛ, Djeoromitxí *-tu ‘powder, flour’.

The classifiers also are found fossilized with nouns, as *-ko in *tʃako ‘mouth’ and *-re in Arikapú nukutore and Djeoromitxí nũt(e)re ‘tongue’. The element -kakʰ ‘shell of a nut, seed or animal’ itself seems to represent a composite of two classifiers. It is possible that the classifiers originally represented compounded nouns, as analyzed by Pires (1992) for Djeoromitxí. They may have been derived from full nouns and represent an intermediate stage in a grammaticalization process. When looking at the other languages of the region (see van der Voort 2005 and Crevels and van der Voort 2008), several genealogically unrelated languages—such as Nambikwara languages, Tacana languages, and the isolated languages of Rondônia—display similar systems. Some even have similar classifying forms, and it is conspicuous that those forms resemble the Jabutí classifiers that resist reconstruction, such as Kwaza (isolate) -mẽ ‘porridge’. Then again, the reconstructed Proto-Jabutí forms *nũ ‘food, porridge’ and *-nɨ ‘thorn’ are also widely attested in the region, respectively as -nũ ‘(fine granular matter such as powder, hair, porridge)’ in various unrelated languages and as -nɨ ‘thorn, needle’ in Kwaza. Especially with respect to Proto-Jabutí *-nɨ, likely cognates occur in several Macro-Jê languages, such as Apinajé (Northern Jê) ni, Rikbaktșá (i-)ni. This suggests that if the Kwaza and Jabutí forms indeed have the same origin, Jabutí was probably the source.

It is important to be aware of grammatical traits that are similar to those in other languages of the region and that could represent areal features. Either such traits may have been transferred to the Jabutí languages through diffusion from other languages, or the similarity may have been an extra motivation for original Proto-Jabutí traits to be preserved in the Jabutí languages. Alternatively, as will be seen in the following section, some of these grammatical elements are reconstructible for Proto-Jê. Thus, the similarity between
such elements in the Jabutí languages and unrelated local languages can also point to (Proto-)Jabutí as a source of regional diffusion.29

3. The Macro-Jê stock. This section briefly discusses the composition of Macro-Jê, providing a succinct evaluation of the evidence offered thus far as support for the hypothesis of genetic relationship among the several members of the stock (3.1) and a note on their geographic distribution and internal classification (3.2). In addition, we discuss the overall characteristics of the languages included in the stock (3.3), thus setting a background against which the comparison between the Jabutí family and well-established Macro-Jê languages can take place (4). The discussion of the main characteristics of Macro-Jê relies heavily on data from the Jê family, the principal member of the stock, and, for obvious historical reasons, one whose inclusion in it was never questioned. This decision—of comparing Jabutí mainly with Jê, instead of the whole stock—has a number of methodological advantages, the main one being that reconstructions of Proto-Jê, the ancestor of the Jê languages, are available (Davis 1966 and Ribeiro [forthcoming]), providing us with a temporal depth30 that is hardly available elsewhere in Macro-Jê (the remaining members of the stock being mostly single-member families; see table 9). Also, a preliminary comparison suggests that Jabutí is not particularly closer genetically to any other families included in the stock, including those that are geographically nearer, such as Boróro, Rikbaktsá, and Chiquitano. Therefore, there is no need for, and probably little gain from, including these languages in the present comparison. In addition, the geographic location of the Jabutí and Jê families (see the map in figure 1) makes it rather unlikely that any similarities detected are due to contact; in fact, no clear cases of shared vocabulary in areas which are more amenable to borrowing (material culture, fauna, flora) have so far been detected. However, data from families for which reliable documentation is available and whose phonological correspondences with Jê are better established, namely, Karajá (Davis 1968 and Ribeiro 2004a), Ofayé (Gudschinsky 1971 and Ribeiro 2005), Maxakalí (Davis 1968), and Krenák (Seki 2002), will be used whenever suitable cognates are found.

29 According to the glottochronological method (as described in, e.g., Jeffers and Lehiste 1992), the time depth between Arikapú and Djeoromitxí could be estimated at about 18 centuries, based on the percentage (58%) of reconstructed shared vocabulary from the Swadesh 100-word list (see Appendix A). We have not adopted this estimate as a fact, however, because the glottochronological method has been proven to be unreliable.

30 Greg Urban estimates the age of Proto-Jê as “three thousand years or more,” suggesting that “[T]he Jê languages are more diversified internally than the ones of the Romance family of Indo-European” (Urban 1998:90 [our translation]).

The Macro-Jê stock comprises the Jê family and a number of possibly related language families, most of which are located in Brazil. The term “Macro-Jê” was introduced by Mason (1950), replacing earlier labels such as “Tapuya” and “Tapuya-Jê” (for a historical overview of the Macro-Jê stock, see Rodrigues 2002). Although there is agreement on the inclusion of most of the families (table 9), recent classifications (Rodrigues 1986, Greenberg 1987, and Kaufman 1990) differ as to the precise scope of Macro-Jê. Given the lack of comprehensive comparative studies, Macro-Jê remains largely a “working hypothesis” (Rodrigues 1999), the present paper being a good example of its tentative nature. Two other examples will suffice: Guató and Chiquitano. In the case of Guató, whose inclusion is agreed upon by all major classifications, the purported evidence is particularly scanty, and a

Fig. 1.—Distribution of the Macro-Jê language families.
thorough search for cognates in the main source on the language (Palácio 1984) reinforces our belief that this language should not be included in the stock. On the other hand, although the inclusion of Chiquitano has been far from a consensus, recent studies (Adelaar 2008 and Santana 2006) have uncovered interesting grammatical, phonological, and lexical similarities, far

---

### TABLE 9
**THE MACRO-JÊ HYPOTHESIS**

<table>
<thead>
<tr>
<th>1. Jê</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Southern Jê</em></td>
</tr>
<tr>
<td>Kaingáng, Xokléng, †Ingaín</td>
</tr>
<tr>
<td><em>Amazonian Jê</em></td>
</tr>
<tr>
<td>Northern Jê: Panará, Suyá (including Tapayúna), Kayapó, Apinajé, Timbíra (Paraktêjê, Krahô, Pykobyê, etc.)</td>
</tr>
<tr>
<td>Central Jê: Xavânte, Xerênte, †Akroá-Mirim, †Xakriabá †Jeikó</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>2. Kamaká</th>
</tr>
</thead>
<tbody>
<tr>
<td>†Kamaká, †Mongoyó, †Menién, †Kotoxó, †Masakará</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>3. Maxakalí</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maxakalí, †Pataxó, †Kapoxó, †Monoxó, †Makoní, †Malalí</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>4. Krenák</th>
</tr>
</thead>
<tbody>
<tr>
<td>Krenák (Botocudo, Borúm)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>5. Purí (Coroado)</th>
</tr>
</thead>
<tbody>
<tr>
<td>†Coroado, †Purí, †Koropó</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>6. Ofayé</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ofayé</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>7. Rikbaktsá</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rikbaktsá (Canoeiro)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>8. Boróro</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boróro, †Umutína, †Otúke</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>9. Karajá</th>
</tr>
</thead>
<tbody>
<tr>
<td>Karajá (including four dialects, Southern Karajá, Northern Karajá, Javaé, and Xambioá)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>10. Karirí</th>
</tr>
</thead>
<tbody>
<tr>
<td>†Kipeá, †Dzubukuá, †Pedra Branca, †Sabuyá (included by Rodrigues, but not Greenberg and Kaufman)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>11. Jabutí</th>
</tr>
</thead>
<tbody>
<tr>
<td>Djeoromitxí (Jabutí), Arikapú (included by Nimuendajú [2000] and Greenberg but not Rodrigues and Kaufman)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>12. Yatê</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yatê</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>13. Guató</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guató</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>14. Chiquitano</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chiquitano (Besiro) (included by Greenberg and Kaufman but not Rodrigues)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>15. Otí</th>
</tr>
</thead>
<tbody>
<tr>
<td>†Otí (Eo-Xavânte) (the inclusion of Otí, proposed only by Greenberg, is not substantiated by the available data)</td>
</tr>
</tbody>
</table>

---

For more on the Macro-Jê Hypothesis, see Ribeiro (2006).
Extinct languages are indicated by †.
more compelling than those proposed as evidence for the inclusion of Guató (Rodrigues 1986; 1999).

The only family-level reconstructions available are Davis (1966) and Ribeiro (forthcoming), for Proto-Jê.\textsuperscript{31} So far, lexical comparative evidence supporting the inclusion of individual families in the Macro-Jê stock has been presented for Kamakã (Loukotka 1932), Maxakalí (Loukotka 1931; 1939 and Davis 1968), Purí (Loukotka 1937), Boróro (Guérios 1939), Krenák (Loukotka 1955 and Seki 2002), Karajá (Davis 1968), Ofayé (Gudschinsky 1971 and Ribeiro 2005), and Rikbaktsá (Boswood 1973). In addition, some studies have shown very suggestive cases of morphological idiosyncracies shared by Jê, Boróro, Maxakalí, Karirí, Karajá, and Ofayé (Rodrigues 2001 and Ribeiro 2002a; 2004a; 2007). In sum, evidence for the inclusion of different families in the Macro-Jê stock is rather uneven, ranging from the fairly proven (Maxakalí, Krenák, and other, extinct Eastern Macro-Jê languages, Ofayé, Karajá, etc.) to the virtually untested (Guató and, to a lesser degree, Yatê). Even in those cases for which inclusion in the Macro-Jê stock is beyond question, much more comparative research is needed so that additional cognates can be identified and more accurate correspondences can be established.\textsuperscript{32} The map in figure 1 shows the distribution of all possible members of the Macro-Jê linguistic stock. The numbers correspond to the language families listed in table 9.

3.2. \textit{Geographic distribution and internal classification.} With the exception of Chiquitano (which is spoken mostly in Bolivia, with a small number of speakers in Mato Grosso, Brazil), all Macro-Jê languages are spoken in Brazilian territory, although in the past Otúke (Boróro) and Ingáín (Southern Jê), both now extinct, were spoken in Bolivia and Argentina, respectively. Although Rikbaktsá and some Jê languages (Kayapó, Suyá, Panará) are spoken on the southern fringes of the Amazon, the overall distribution of Macro-Jê languages is typically non-Amazonian. Yatê, Krenák, and Maxakalí languages are spoken in eastern Brazil, the same having been the case for Purí, Kamakã, and Karirí (all now extinct). Since several pur-

\textsuperscript{31} For recent attempts at reconstructing Proto-Purí and Proto-Kamakã, see Silva Neto (2007) and Martíns (2007), respectively.

\textsuperscript{32} Thus, the ongoing study of Ofayé by one of the authors (Ribeiro) makes possible an assessment of the previous work done by Gudschinsky (1971; 1974)—until recently, the main source available on this severely endangered Macro-Jê language. As it turns out, Gudschinsky’s comparative work presents serious methodological problems, including arbitrary word segmentations and the failure to recognize borrowings (Ribeiro 2005). Also, for Rikbaktsá, Boswood’s estimate of 38% cognacy seems to be too optimistic, since many of her purported correspondences are very poorly attested (Boswood 1973). If that is the situation for families for which comparative studies were conducted, one can assume that the situation is far less satisfactory for the remaining ones.
ported Macro-Jê languages were spoken in eastern Brazil, a number of them became extinct early on, under the impact of European colonization, Yatê being the only surviving indigenous language in the Brazilian northeast (not including Maranhão). Central and Northern Jê tribes (here included under a single branch of the Jê family, Amazonian Jê), as well as the Boróro and the Ofayé, traditionally occupy the savanna areas of Central Brazil. The southernmost Macro-Jê languages are those belonging to the southern branch of the Jê family, spreading from São Paulo to Rio Grande do Sul. Karajá is spoken along the Araguaia River, in Central Brazil. The traditional Guató territory is the Paraguay River, near the Bolivian border.

On geographic grounds, Rodrigues (1999) divides his Macro-Jê families into three groups: Eastern (Purí, Krenák, Maxakalí, Kamakã, Karirí, and Yatê), Central (Karajá and Jê), and Western (Boróro, Guató, Ofayé, and Rikbaktsá). However, he does not attempt to provide internal classifications on genetic grounds. A first attempt at internal genetic classification was recently made by Ribeiro (2007), who suggests that four out of Rodrigues’s six eastern families (Maxakalí, Krenák, Kamakã, and Purí) may form a subgroup inside the stock. If that claim is confirmed by further research, this will have important implications for theories about the original homeland of Proto-Macro-Jê, and for claims such as the one made in this paper (see 5).

3.2.1. A note on Proto-Jê. The existence of Jê as a language family has been recognized since early classifications of Brazilian languages (von Martius 1867), although the inclusion of Kaingáng (and closely related Southern Jê languages) was traditionally controversial. For instance, Loukotka (1932) initially included Kaingáng in the Jê family but later on changed his mind (Loukotka 1939). On the other hand, Nimuendajú, in his late years, was convinced that Kaingáng was indeed part of the family (Nimuendajú and Guérios 1948). This issue remained unsettled until Davis demonstrated, on comparative grounds, that Kaingáng should be included in the family, although admitting it to be “a rather divergent member of the family” (Davis 1966:11). As the contentious history of the inclusion of Kaingáng suggests, both Northern and Central Jê languages are much closer to each other than to Southern Jê languages, forming a particular branch (here termed “Amazonian Jê”) inside the family. This has obvious implications for Macro-Jê studies, since a form attested only for a branch of the family may very well be an innovation. Davis did not take into account this potential problem; about one-third of his reconstructions were attested only for Amazonian Jê. Furthermore, Davis’s comparative work presents a number of additional problems, including false cognates, lack of morphological information, and incorrect reconstructions.

A case in point is the absence, in Davis’s reconstruction, of word-initial *w. Although Northern Jê languages are conservative in several aspects (for
instance, the preservation of essentially the entire vowel inventory of Proto-
Jê), the fact that many of Davis’s Proto-Jê forms are identical with Apinajé
words may mean that Davis relied too much on data from this language (at
that time one of the best-described Jê languages). This had a number of un-
derirable consequences for his reconstruction, since Apinajé presents signif-
icant innovations in the consonantal system. For instance, Davis reconstructs
the roots meaning ‘to kill’ and ‘wood’ as homophonous. However, as the
examples in table 10 show, they are only homophonous in Apinajé, since
Apinajé merged PJê *w and *p.

Given these and other shortcomings of Davis’s reconstruction, the Proto-
Jê (and Proto-Amazonian Jê) forms adopted in this paper are exclusively
those reconstructed by Ribeiro (forthcoming); additional comparative remarks
on Proto-Jê and its relationship to other families are from the same source.

3.3. Main characteristics of Macro-Jê. Macro-Jê languages are typo-
logically very similar, a fact which, in and of itself, cannot be taken as a sign
of genetic relationship. The majority of Macro-Jê languages are verb-final,
with postpositions instead of prepositions and possessor-possessed order in
genitive constructions—exceptions being Guató (which, as suggested above,
should probably be excluded from the stock), Chiquitano, and Karirí. For the
latter two, however, internal evidence seems to point to an older SOV pat-
tern, revealed by the behavior of adpositions and compounds (Rodrigues
1999 and Ribeiro 2007, for Karirí; Adelaar 2008, for Chiquitano).

Most Macro-Jê languages have a relatively simple morphology. In most
languages, including those of the Karirí, Krenák, Jê, Ofayé, and Maxakalí
families, productive inflectional morphology is limited to person marking,
the same paradigms being generally shared by nouns, verbs, and adpositions
alike. Table 11 illustrates the series of singular personal prefixes in Jê lan-
guages, all of which can be reconstructed for Proto-Jê, many with likely cog-
nates in other Macro-Jê families, such as Maxakalí and Karajá.

The personal prefix series includes a linking morpheme (traditionally called
“relational prefix” [Rodrigues 2001]), which occurs with most vowel-initial
stems whenever they are immediately preceded by an absolutive argument.

<table>
<thead>
<tr>
<th>Davis</th>
<th>Apinajé</th>
<th>Suyá</th>
<th>Kayapó</th>
<th>Xerénte</th>
<th>Kaingáng</th>
<th>Ribeiro</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘wood, firewood’</td>
<td>*pí</td>
<td>pí</td>
<td>*wí</td>
<td>pí</td>
<td>mmí</td>
<td>pí</td>
</tr>
<tr>
<td>‘to kill’</td>
<td>*pí</td>
<td>pí</td>
<td>pí</td>
<td>bí</td>
<td>wí</td>
<td>(PAJê only)</td>
</tr>
</tbody>
</table>
the jabutí language family

(possessors, with nouns; objects, with transitive verbs and postpositions; and, in some languages, subjects, with intransitive verbs). As the Suyá (44) and Karajá (45) examples below illustrate, most vowel-initial roots present at least two stem-forms, one with the linking prefix (cognate with Proto-Jê *j-*) and the other with a third-person marker (cognate with Proto-Jê *s-*):

(44a) *katireye y-aykwa
    boy lp-mouth
    ‘the boy’s mouth’

(44b) s-aykwa
    3-mouth
    ‘his mouth’

(45a) *dore d-e
    parrot lp-wing
    ‘the parrot’s wing’

(45b) d-e
    3-wing
    ‘its wing’

Notice that two third-person markers are reconstructed for Proto-Jê: *i-* occurs with consonant-initial stems, whereas *s-* occurs with vowel-initial stems (in complementary distribution with the linking prefix *j-*). For the sake of clarity, j- stems (that is, stems which present *s-/*j- alternating forms) are given here preceded by the linking prefix. Considering that the equivalents of *j-* in other families (as well as in Kaingáng, a Southern Jê language) may be reanalyzed as a part of the stem, it seems appropriate to

In a few cases, it is the third-person marker *s-* which becomes fossilized as part of the root.

---

TABLE 11
PERSONAL PREFIXES IN JÊ

<table>
<thead>
<tr>
<th></th>
<th>NORTHERN JÊ</th>
<th>CENTRAL JÊ</th>
<th>SOUTHERN JÊ</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Proto-Jê</td>
<td>Apinajé</td>
<td>Parkatêjê</td>
</tr>
<tr>
<td>1</td>
<td><em>i-</em></td>
<td>i-*</td>
<td>i-*</td>
</tr>
<tr>
<td>2</td>
<td>*a-</td>
<td>a-</td>
<td>i-*</td>
</tr>
<tr>
<td>3</td>
<td>*i-</td>
<td>i-</td>
<td>0-</td>
</tr>
<tr>
<td>3</td>
<td>*s-</td>
<td>0-</td>
<td>h-</td>
</tr>
<tr>
<td>REL</td>
<td>*j-</td>
<td>j-, tf, n</td>
<td>j-, tf-</td>
</tr>
</tbody>
</table>

---

33 In a few cases, it is the third-person marker *s-* which becomes fossilized as part of the root.
use stems (linking prefix + root), rather than bare roots, as the basis for the comparison.

In addition to personal prefixes (including linking morphemes), Jê languages also present a number of semi-fossilized prefixed elements (which, as the personal prefixes, occur with both noun and verb stems), whose semantic and grammatical properties tend to be rather vague. As Oliveira describes them, for Apinajé, such elements are “fossilized morphological debris that simply subcategorize words into morphological or semantic classes at the synchronic level” and “tend to be semantically empty or of elusive meaning” (2005:82). Most of these morphemes (a sample of which includes Apinajé č-u-, j-ar-, ku-, and ka-) can be reconstructed for Proto-Jê. The examples in (46), from Apinajé, illustrate the classificatory value of some of these morphemes:

\[(46a)\] kuʔo ‘wash hard objects’
\[(46b)\] kaʔo ‘wash soft objects’

Although they tend to be relatively fossilized, they were probably still productive in the proto-language, since cognate stems may take different prefixes in different languages (or may not take any prefix at all). Compare the following examples from Apinajé (Northern Jê) and Kaingáng (Southern Jê):

\[(47a)\] PJê *mi ‘tail’ > Apinajé j-ami vs. Kaingáng mi
\[(47b)\] PJê *pri ‘path’ > Apinajé pri vs. Kaingáng jāpri
\[(48a)\] PJê *wē ‘to speak’ > Apinajé kapē, Kaingáng wī
\[(48b)\] PJê *tēm ‘to come’ > Apinajé tēm, Kaingáng kātīŋ

Tense and aspect distinctions are generally conveyed by particles and auxiliaries, rather than inflectionally. Noun incorporation is rare, having been reported for Karajá and a few Northern Jê languages such as Panará and Parkatêjê. There are, however, a few fossilized noun–verb compounds, which may have resulted diachronically from noun incorporation (e.g., Apinajé j-arkje ‘to yawn’ < PJê *j-ar- ‘mouth’ + *kje ‘open’; Kaingáng nipe ‘to wash hands < PJê *j-ī- ‘hand’ + pe ‘to wash’; cf. Maxakalí yīm ‘hand’; Ofayé j-ī ‘hand’; Karajá d-ē- ‘arm’).

A few languages present more complex morphologies, but in those cases for which comparative evidence is available, morphological complexity can be shown to be an innovation. Panará is a case in point, providing an interesting example of “polysynthesis in the making” (when compared to closely related Northern Jê languages), with productive noun incorporation, applicatives (originated from incorporated postpositions), and classifiers (Dourado 2001; 2002). Karajá also presents a fairly complex verb morphology, but, once again, the “core” morphology, comprising derivational suffixes and pronominal prefixes immediately attached to the stem, is what has clear cog-
nates in Macro-Jê (as would be expected; the closer to the stem, the more likely it is for an affix to be older [see Givon 2000]).

While inflectional morphology is predominantly prefixing, derivational morphology is generally suffixing. As with inflectional prefixes (table 11), most derivational suffixes in Jê are retentions from Proto-Jê (or, at least, from Proto-Amazonian Jê). These include a causativizer (PJê *-n), diminutive (PAJê *-re), augmentative (PAJê *-ti), an agent-noun suffix (PAJê *-jua), an instrument-noun suffix (PJê *-j3), and an abstract nominalizer (PAJê *-r). The latter three have likely cognates in at least another Macro-Jê language, Karajá (-du ‘subject-noun suffix’, -da ‘instrument-noun suffix’, -r- ‘nominalizer’ [see Ribeiro 2008]).

When compared to languages of other lowland South American families (such as Carib and Tupí), Jê languages typically present larger vowel inventories. As table 12 illustrates, Proto-Jê had a system of ten oral vowels and at least four nasal ones. Karajá has an even larger vowel inventory (with twelve oral vowels and three nasal ones), but most Macro-Jê languages display smaller inventories; Maxakalí, for instance, has only five oral vowels, all with nasal equivalents (Popovich and Popovich 2005). As comparative

34 The examples below illustrate the categories for which the Karajá verb inflects: person, direction, voice, and object (tense is expressed by a clitic). Notice that the prefix d- ‘third person’ (i) is a cognate of Proto-Jê *s- and, as in Jê, occurs with noun stems as well. The verb stem, l-adj, is cognate with Proto-Jê *j-i ‘to lay’ and presents in both families an alternation involving the relational prefix: Karajá l-adj (ii) vs. d-adj (i); PJê *j-i vs. *s-i. Example (iii) illustrates another Karajá verb stem with a Jê cognate (rt ‘to leave behind’, PJê *re), taking the second-person object prefix a- (again, a cognate with Proto-Jê *a-).

(i) ka-r-i-d-adj=kre
1-CTRG-TRANS-3-leave=FUT
‘I will lay it down’.

(ii) Ø-r-t-wa-l-adj=kre
1-CTRG-TRANS-1-REL-leave=FUT
‘She will lay me down’.

(iii) ka-r-Ø-a-rt=kre
1-CTFG-TRANS-2-leave=FUT
‘I will leave you’.

<table>
<thead>
<tr>
<th>Oral</th>
<th>Nasal</th>
</tr>
</thead>
<tbody>
<tr>
<td>i</td>
<td>i</td>
</tr>
<tr>
<td>t</td>
<td>(i)</td>
</tr>
<tr>
<td>e</td>
<td>ø</td>
</tr>
<tr>
<td>ø</td>
<td>t̂</td>
</tr>
<tr>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td>ê</td>
<td>ê</td>
</tr>
<tr>
<td>ń</td>
<td>ñ</td>
</tr>
<tr>
<td>u</td>
<td>ũ</td>
</tr>
<tr>
<td>ĩ</td>
<td>ĩ</td>
</tr>
<tr>
<td>ễ</td>
<td>ễ</td>
</tr>
<tr>
<td>ẽ</td>
<td>ẽ</td>
</tr>
<tr>
<td>ĩ</td>
<td>ĩ</td>
</tr>
<tr>
<td>ũ</td>
<td>ũ</td>
</tr>
<tr>
<td>ñ</td>
<td>ñ</td>
</tr>
<tr>
<td>t̂</td>
<td>t̂</td>
</tr>
<tr>
<td>ễ</td>
<td>ễ</td>
</tr>
<tr>
<td>ẽ</td>
<td>ẽ</td>
</tr>
<tr>
<td>ĩ</td>
<td>ĩ</td>
</tr>
<tr>
<td>ũ</td>
<td>ũ</td>
</tr>
<tr>
<td>ñ</td>
<td>ñ</td>
</tr>
<tr>
<td>t̂</td>
<td>t̂</td>
</tr>
<tr>
<td>ễ</td>
<td>ễ</td>
</tr>
<tr>
<td>ẽ</td>
<td>ẽ</td>
</tr>
<tr>
<td>ĩ</td>
<td>ĩ</td>
</tr>
<tr>
<td>ũ</td>
<td>ũ</td>
</tr>
<tr>
<td>ñ</td>
<td>ñ</td>
</tr>
<tr>
<td>t̂</td>
<td>t̂</td>
</tr>
<tr>
<td>ễ</td>
<td>ễ</td>
</tr>
<tr>
<td>ẽ</td>
<td>ẽ</td>
</tr>
<tr>
<td>ĩ</td>
<td>ĩ</td>
</tr>
<tr>
<td>ũ</td>
<td>ũ</td>
</tr>
<tr>
<td>ñ</td>
<td>ñ</td>
</tr>
</tbody>
</table>
evidence becomes available, it seems clear that Jê is indeed more conservative, whereas languages such as Maxakalí underwent considerable merging, leading to smaller vowel inventories.

Northern Jê languages inherited the Proto-Jê vocalic system practically unchanged (as illustrated by Suyá [Santos 1997]), whereas Central and Southern Jê languages underwent pervasive processes of vowel shift—although the end results, in terms of inventory, are vocalic systems which are only minor deviations from Proto-Jê (major phonological contrasts, such as the distinction between open and close-mid vowels, having been preserved). Table 13 illustrates the process of vowel shift involving front vowels in Central Jê, with systematic raising of *e to /e/ and of *i to /i/; the gap that would have resulted from the raising of *e is filled by the fronting of *s.

Another common phenomenon in Jê languages (as well as in Ofayé) is the insertion of “echo vowels” at the end of words ending in a consonant (most commonly, approximants): OFY ʰar ‘foot’ [ʰara’], etc. Such extraprosodic vowels are generally unstressed, but in (Proto-)Central Jê they seem to have been reanalyzed as part of the stem. In Xerente, where there is a strong tendency toward vowel deletion (both diachronically and synchronically), the original vowel (probably, after losing stress) is deleted: *[para] ‘foot’ > *[pa’ra] > [pra]. Comparatively, this process is important because it provides an explanation for apparent (and illusory) cases of metathesis: PJê *par > XER pra. The fact that such correspondences are found inside the Jê family makes a comparison such as PJê *par :: PJab *praj more plausible. A similar process was probably at play in Karajá, which would explain why a correspondence involving a (C)CVC word in Proto-Jê has the vowel showing up in a different syllable in Karajá:

(49a) PJê *j-am, KRI l-əba ‘to stand’
(49b) PJê *pram, KRI rəba ‘hunger’
(49c) API mən, KRI bədə ‘honey’
(49d) API mur, KRI bəro ‘to cry’

Syllabic patterns are rather simple, obstruent clusters being limited to C₁C₂V (where C₁ is a plosive and C₂ is *r): PJê *pri ‘path’, *kra ‘offspring’, *pram ‘hunger’, etc. Proto-Jê, as well as the present Jê languages, allows
syllables to end in a consonant, including stops. In Kaingáng, Proto-Jê final stops underwent systematic lenition (PJê *pek ‘fart’ > KAI pej; PJê *tpt ‘hard’ > KAI tar, etc.). Stress is generally predictable, falling on the last syllable, with the apparent exception of Rikbaktsá (Silva 2005). Phonologically contrastive tone oppositions occur in Yatê and Guató (Palácio 1984). Processes such as nasal spreading and vowel harmony are generally absent. An exception is Karajá, which presents [ATR] (“advanced tongue root”) vowel harmony, a rare phenomenon among South American languages (Ribeiro 2002b).

As table 14 shows, there was no contrast between voiced oral and nasal plosives in Proto-Jê (a situation that is preserved in most Jê languages). This also occurs in several other Macro-Jê families, including Karajá, Maxakalí, Krenák, and Ofayé. When one considers the reflexes in all the branches of the family (as well as correspondences with other families), it is likely that the nasal consonants were pronounced as fully nasal around nasal vowels and as semi-nasalized voiced stops elsewhere.35

Although a few aspects of the Proto-Jê phonological inventory remain to be worked out,36 phonological correspondences between the different members of the family are mostly well established. That is obviously important for comparative purposes, since it enables us to detect likely cognates even

35 The symbol *s in table 14 is used to represent a phoneme whose regular reflexes are /s/ in Panará, /s/ in Suyá, /h/ in Timbira, /s/ in Central Jê (/h/, before PJê *i; /z/ before PJê *i), zero (or a glottal stop) in Apinajé and Kayapó, /j/ in Kaingáng, and /ð/ in Xokléng. Davis (1966) reconstructs it as *z, but when one considers the reflexes listed above, *s seems to be a better approximation to its likely pronunciation.

36 These are indicated by the parentheses in the inventory tables. Most Jê languages present glottal consonants (a stop, a fricative, or both), but comparative evidence for their reconstruction is still inconclusive (except for Timbira /h/ and Apinajé /l/, which tend to be regular reflexes of Proto-Jê *s, and instances of Central-Jê /h/, which trace back to *k, before Proto-Jê high vowels). Possible cognates found so far are still questionable. For instance, the existence of Kaingáng ho ‘body’ and Central Jê ho ‘body’ would in principle allow the reconstruction of a form with a glottal fricative in Proto-Jê (*ho ‘body’), but the vowel correspondence would be irregular (Kaingáng *a/ corresponds regularly to Central Jê /a/); in addition, the Central Jê form could be related to ho ‘skin’ (< PAJê *ki ‘skin’). Therefore, the similarity between Kaingáng and Central Jê in this case could be accidental. The reconstruction of a glottal stop seems at first to be corroborated by the existence of Kaingáng pâ?i ‘chief’ (Xokléng pa?i) and Apinajé pa?i ‘chief’ (cognates of
in cases for which a word cannot (so far) be reconstructed for one (or both) of the proto-languages (see 4.1).37

4. The evidence for the Jabutí—Macro-Jê hypothesis. As mentioned in the introductory sections, the Jabutí languages used to be regarded by some as being related to the Jê languages. Nimuendajú includes them in the Jê linguistic family on his ethno-historical map (IBGE 1981 [1944]) on the basis of lexical data provided by Snethlage, and Greenberg (1987) classifies them as belonging to the Macro-Jê stock. This classification is usually accepted by typological linguists, but not by the majority of Americanists. In most other authoritative works on South American languages, such as Rodrigues (1986), Campbell (1997), Dixon and Aikhenvald (1999), Loukotka (1968), and Kaufman (1990), the Jabutí languages are assumed to constitute an isolated family. Now that abundant data of high quality have finally been obtained, it is time to revise the classification of Jabutí languages and to evaluate the possible consequences this may have for current ideas about the prehistoric expanse of the speakers of Macro-Jê languages.38 In the follow-

which occur in other Northern Jê languages as well), leading to the reconstruction of *paʔi for Proto-Jê. There is a possibility, however, that this may be a loanword from Tupí-Guarani languages, possibly introduced independently in Southern Jê and Northern Jê (cf. Guarani paʔi ‘priest; head of an extended family’, etc.). As for the vowels, the existence of the nasals *u and *i, reconstructed by Davis (1966), is also questionable (perhaps even synchronically, in the Northern Jê languages for which they are described as being phonemic). The only item that Davis reconstructs as containing the vowel *i is ‘to sit’, reconstructed by Davis as *ji. However, in both Central and Southern Jê, the vowel in this stem is the same as obvious reflexes of Proto-Jê *u. Since in Northern Jê languages, [ɨ] has an extremely limited distribution (occurring generally around glides), it is likely that this pronunciation emerged as the result of height assimilation. Whether this analysis is synchronically valid as well is a subject for further research. Similar arguments can be used against the reconstruction of *u.

37 This is a methodological point that needs to be stressed, especially inside the field of Macro-Jê comparative studies, where claims of genetic relationship are often based on superficial similarities. A case in point is Rodrigues’s purported evidence for a genetic relationship between (Macro-)Jê and Tupí, a major part of which involves an apparent correspondence between Proto-Tupí-Guarani *p and Kaingáng /h/ (e.g., Proto-Tupí *pem ‘to weave’, KAI fe ‘to spin’, etc. [Rodrigues 1985:395–96; 2000:102]). However, a comparative look at other (Macro-)Jê languages shows that Kaingáng /h/ traces back to Proto-Jê *s: PJê *se ‘thread’ > Panará se, Parkatêjê he, Xokléng de, etc. Kaingáng fe is actually a derived form, including the noun root fe ‘thread’ and the causativizing suffix -n (< PJê *-n; vowel lowering is regularly triggered by the suffixation of -n). Cases such as these make one suspect that many of the purported Kaingáng/Tupí cognates proposed by Rodrigues are accidental, superficial similarities (for further details, see Ribeiro [forthcoming]).

ing sections, we present lexical and grammatical evidence for the genetic relationship between the Jê and Jabutí families (corroborated, whenever possible, by data from other Macro-Jê families).

4.1. Lexical evidence. In this section, we present lexical comparative evidence for the genetic relationship between the Jabutí and Jê families, based on recent reconstructions of both Proto-Jabutí, by van der Voort (2007), and Proto-Jê, by Ribeiro (forthcoming). Note that both reconstructions took place independently from one another: van der Voort reconstructed Proto-Jabutí without any prior knowledge of (Macro-)Jê linguistics, while Ribeiro had over the years produced comparative studies of Macro-Jê languages without ever having looked at Jabutí. Only after the basic facts of both proto-languages were established were the results compared.

The likely cognates are distributed into three different tables, according to their degree of attestation. Table 15 includes only elements which can be reconstructed for both proto-languages. Table 16 includes items that can be reconstructed for only one of the proto-languages, as they are attested only in individual languages in the other family. Finally, table 17 includes items that were attested only in individual members of both families. As much as possible, likely cognates which would require arbitrary “morphological” segmentations were avoided.39 In all three tables, the first column includes information on whether an item is part of Swadesh’s 100-word list; the numbers refer to the original numbers in Swadesh’s list (see Appendix A).

Most correspondences are rather straightforward. As the descriptive summaries given above show, Jabutí and Jê (as Macro-Jê languages in general) are isolating, presenting very little productive morphology. The comparative study suggests that this is an inherited feature of these languages and not a result of independent phonological or morphological erosion. Most reconstructed stems, in both proto-languages, are monosyllabic or (though rarely) disyllabic; the few exceptions are old compounds (e.g., PJê *j-akua ‘mouth’,40 PJab *tʃako) or possible loans.41

39 The few exceptions—three, to be exact—are cognates 18, 35, and 46. For the possible existence of a morphological entity *tfu- in Jabutí (cf. cognate 35), see Dje runa vs. Ari nô ‘to sit’. Further research on the Jabutí languages may help corroborate this hypothesis.

40 For Proto-Jê, Ribeiro (forthcoming) reconstructs two items with the meaning ‘mouth’: *j-ar-, an element that occurs in compounds (cf. PRk j-ar-ko ‘saliva’ < *j-ar + *yo ‘water’), and a freestanding word *j-arqua (probably an original compound including *j-ar- and a so far unidentified morpheme). Two similar forms may also be reconstructed for Proto-Jabutí: a freestanding word, *tʃako (which probably includes the morpheme *ko ‘hole’) and a compound element *tʃa (cf. Ari tʃat ‘labret; lit., mouth-ornament’ vs. tʃat ‘upper-arm bracelet; lit., arm-ornament’).

41 An example is the word for ‘anteater’, reconstructed for Proto-Jabutí as *patʃuri, a form that is strikingly similar to Karajá wariri ‘anteater’ (also a possible cognate is the Amazonian Jê form, e.g., Apinajé pat). Although the phonological correspondences seem to be regular, similar forms occur in Karib and Tupí languages as well. Therefore, even though this may
Despite its cursory nature, the sample of likely cognates provided in tables 15–17 shows many regular correspondences:

PJê *p :: P Jab *p (5, 18)
PJê *r :: P Jab *t (7, 17)

eventually turn out to be a legitimate cognate, it cannot, at this stage of the research, be used as a piece of evidence for genetic relationship.
### TABLE 16
**Jê/Jabuti Cognates II: In Actual Languages and Proto-Languages** (Swadesh)

<table>
<thead>
<tr>
<th>(Swadesh)</th>
<th>Jê</th>
<th>Jabutí</th>
<th>Other Families</th>
</tr>
</thead>
<tbody>
<tr>
<td>24 (26) *j-are</td>
<td>*j-are</td>
<td>dje rari</td>
<td>‘root’</td>
</tr>
<tr>
<td>25 (8)</td>
<td>*mā</td>
<td>*mā</td>
<td>‘no’</td>
</tr>
<tr>
<td>26</td>
<td>*mā</td>
<td>dje mā</td>
<td>‘dative’</td>
</tr>
<tr>
<td>27 (36)</td>
<td>*j-ak</td>
<td>dje rari</td>
<td>‘wing’</td>
</tr>
<tr>
<td>29</td>
<td>API mop</td>
<td>*mu</td>
<td>‘yam’</td>
</tr>
<tr>
<td>30</td>
<td>API mē</td>
<td>dje me</td>
<td>‘honey’</td>
</tr>
<tr>
<td>31</td>
<td>API *ndj</td>
<td>API *ndj</td>
<td>‘egg’</td>
</tr>
<tr>
<td>32 (33)</td>
<td>*ŋe</td>
<td>API *tē</td>
<td>‘to dance’</td>
</tr>
<tr>
<td>33</td>
<td>API *tē</td>
<td>API *tē</td>
<td>‘to sing; to dance’</td>
</tr>
<tr>
<td>34</td>
<td>API tf-wa</td>
<td>*tpo</td>
<td>‘to bathe’</td>
</tr>
<tr>
<td>35</td>
<td>SUY pi</td>
<td>*jwi</td>
<td>‘go up’</td>
</tr>
<tr>
<td>36 (44)</td>
<td>*ŋ-ō</td>
<td>API nūt</td>
<td>‘tongue’</td>
</tr>
<tr>
<td>37 (16a)</td>
<td>PRK tfe</td>
<td>*dji</td>
<td>‘mother’</td>
</tr>
<tr>
<td>38 (83)</td>
<td>*mē</td>
<td>API mē</td>
<td>‘ashes, dust’</td>
</tr>
<tr>
<td>39</td>
<td>*wi</td>
<td>dje mō</td>
<td>‘to catch’</td>
</tr>
<tr>
<td>40</td>
<td>API *nō</td>
<td>*ū</td>
<td>‘to give’</td>
</tr>
<tr>
<td>41</td>
<td>API mēr</td>
<td>*mō</td>
<td>‘to cry’</td>
</tr>
<tr>
<td>42</td>
<td>API ton</td>
<td>*tōw</td>
<td>‘armadillo’</td>
</tr>
<tr>
<td>43</td>
<td>XAV *nō</td>
<td>*nūrō</td>
<td>‘cord, rope’</td>
</tr>
</tbody>
</table>

PJê *k :: PJab *k (4, 6, 14, 15, 16)
PJê *m :: PJab *m (12, 13)
PJê *j :: PJab *tj (before oral vowels: 6, 9, 10, 11)
PJê *j :: PJab *tj (before front high oral vowel: 23)
PJê *j :: PJab *n (before nasal vowels: 7, 8, 21)
PJê *r :: PJab *r (4, 5, 6)
PJê *s :: PJab *∅ (before PJab high vowels: 19, 20)
PJê *s :: PJab *h (elsewhere)

Having worked out the major correspondences between Proto-Jê and Proto-Jabutí, we can now take into account the cognates in tables 16 and 42 and

---

42 The Karajá form in item 24 occurs in a likely compound, *l-adjkura* ‘cassava’ (cf. *kura* ‘white’).
Besides corroborating and refining the correspondences detected in Table 15, the additional data reveal new correspondences (for instance, PJê *w :: PJab *w). Tables 18 and 19 summarize such correspondences.

As described in the introductory sections, Jê languages present consonant-final syllables, whereas Jabutí languages do not (except for the glide *j). As the comparative data above demonstrate, coda consonants seem to have been lost in Jabutí, with very few exceptions (cognates 7 and 27). Some correspondences are either poorly attested or remain unaccounted for. For instance, examples showing the Proto-Jabutí equivalent of Proto-Jê *n are conspicuously absent in the compared corpora. Correspondences involving Proto-Jê *s are also poorly attested; however, if it turns out that the word-initial consonant alternations found in Djeoromitxí are historically related to similar processes in Jê and other families, then the initial consonant in the morphologically free forms (DJE hapa ‘arm’, hak ‘mouth’, etc.) would be an

43 The Djeoromitxí form in item 51 was documented by Pires (1992) and does not occur in van der Voort’s database.

44 Northern Jê languages are very closely related and lexically conservative. Therefore, all the examples from Apinajé, Parkatêjê, and Suyá provided here are common to the entire Northern Jê branch. Apinajé generally represents well the languages of the branch, but not always. In the case of cognate 35, for instance, Apinajé has a form identical to Suyá: pi. However, when one considers that Apinajé /p/ traces back to both PJê *p and *w, comparing PJab *fjwí to the Apinajé form would be less reliable. Suyá /p/, however, clearly traces back to PJê *w; thus, the Suyá form corroborates the identification of a correspondence between PJê *w and PJab *w.
the Jabutí language family

obvious cognate of the Proto-Jê third-person prefix *s- (cf. Suyá s-ajkwa ‘his mouth’, s-wa ‘his tooth’, etc.).

Both Proto-Jê and Proto-Jabutí have remarkably similar consonantal inventories. The lack of contrast between voiced stops and their nasal counterparts (common in many—if not most—Macro-Jê languages, such as those of the Jê, Maxakalí, and Karajá families) is also preserved in Jabutí. One of the few differences is the absence, in Proto-Jabutí, of a nasal velar *ŋ, a phoneme that tends to be fairly less stable cross-linguistically.

<table>
<thead>
<tr>
<th>Jê</th>
<th>Jabutí</th>
<th>Cognates</th>
</tr>
</thead>
<tbody>
<tr>
<td>*p</td>
<td>*p</td>
<td>5, 18, 31</td>
</tr>
<tr>
<td>*t</td>
<td>*t</td>
<td>7, 17, 36, 42, 52</td>
</tr>
<tr>
<td>*k</td>
<td>*k</td>
<td>4, 6, 14, 15, 16, 46</td>
</tr>
<tr>
<td>*m</td>
<td>*m</td>
<td>12, 13, 25, 26, 28, 29, 38, 44</td>
</tr>
<tr>
<td>*n</td>
<td>*n</td>
<td></td>
</tr>
<tr>
<td>*ŋ</td>
<td>*Ø</td>
<td>32, 22</td>
</tr>
<tr>
<td>*r</td>
<td>*r</td>
<td>4, 5, 6, 24, 27, 32, 33, 36, 38, 46, 47</td>
</tr>
<tr>
<td>*s</td>
<td>*Ø (before PJab high vowels)</td>
<td>19, 20</td>
</tr>
<tr>
<td>*s</td>
<td>*h (elsewhere)</td>
<td>22, 51</td>
</tr>
<tr>
<td>*j</td>
<td>*t (before oral vowels)</td>
<td>6, 9, 10, 11, 24, 27, 47</td>
</tr>
<tr>
<td>*j</td>
<td>*n (before nasal vowels)</td>
<td>7, 8, 21, 30, 36, 43</td>
</tr>
<tr>
<td>*w</td>
<td>*W</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Jê</th>
<th>Jabutí</th>
<th>Cognates</th>
</tr>
</thead>
<tbody>
<tr>
<td>*i</td>
<td>*i</td>
<td>20, 23, 35</td>
</tr>
<tr>
<td>*i</td>
<td>*a</td>
<td>22 (?), 39</td>
</tr>
<tr>
<td>*u</td>
<td>*u</td>
<td>9, 11, 16</td>
</tr>
<tr>
<td>*e</td>
<td>*i</td>
<td>24, 31, 37</td>
</tr>
<tr>
<td>*a</td>
<td>*a</td>
<td></td>
</tr>
<tr>
<td>*o</td>
<td>*u</td>
<td>14, 19, 28</td>
</tr>
<tr>
<td>*e</td>
<td>*e</td>
<td>32, 33, 34</td>
</tr>
<tr>
<td>*ɑ</td>
<td>ARI /i/</td>
<td>47</td>
</tr>
<tr>
<td>*ɛ</td>
<td>DHE /e/</td>
<td>36</td>
</tr>
<tr>
<td>*a</td>
<td>*a</td>
<td>2, 4, 5, 6, 18, 24, 27, 47, 52</td>
</tr>
<tr>
<td>*e</td>
<td></td>
<td></td>
</tr>
<tr>
<td>*ʊ</td>
<td>*ʊ</td>
<td>7, 30, 36, 40, 43, 51</td>
</tr>
<tr>
<td>*ɪ</td>
<td>*ɪ</td>
<td>8</td>
</tr>
<tr>
<td>*ɑ</td>
<td>*ɑ</td>
<td>21, 25, 28</td>
</tr>
<tr>
<td>*ua</td>
<td>*o</td>
<td>6, 10, 34</td>
</tr>
</tbody>
</table>

TABLE 18

Phonological Correspondences between Jê and Jabutí: Consonants

TABLE 19

Phonological Correspondences between Jê and Jabutí: Vowels
Vowel correspondences are less obvious—but, nonetheless, mostly regular. If one assumes (as it seems) that Jê is more conservative, a general process of vowel raising, involving mid vowels as a natural class, seems to have taken place in Pre-Proto-Jabutí, close-mid vowels in Proto-Jê corresponding systematically to high vowels in Jabutí (cf. 24, 31, 37, 14, 19, 28); the data suggest that, as in Central-Jê, close-mid vowels would have merged with the original high vowels in Jabutí. The Proto-Jê sequence *ua is regularly reflected in Proto-Jabutí as a single vowel, *o, possibly as the result of monophthongization (cognates 6, 10, 34). Some gaps in the comparative tables are probably a consequence of the rarity of certain phonemes; Proto-Jê *a, for instance, is particularly rare in the comparative corpus. Additional research will, we hope, help to fill such gaps.

4.1.1. Homophonous pairs. If, as the comparative method assumes, sound changes tend to be regular, and two different words happen to sound the same in a given proto-language, one expects that these words will still be homophonous in the daughter languages, unless lexical replacement or some irregular factor is at play. Consequently, corroborating the regularity of the phonological correspondences postulated above is the fact that cases of homophony in one language tend to correspond to homophonous pairs in the other. Thus, the words for ‘to hear’ (50a) and ‘liver’ (50b) are homophonous in both Proto-Jê (*ma) and Proto-Jabutí (*ma); the words for ‘tooth’ (50c) and ‘to bathe’ (50d), homophonous in Proto-Jabutí (*tfo), are also homophonous in Northern Jê (e.g., Apinajé tf-wa ‘tooth’, tf-wa ‘bathe’; Northern Jê tf, in such cases, traces back to Proto-Jê *j). The words for ‘egg’ (50e) and ‘dance’ (50f), homophonous in Proto-Jê (*yr), are also homophonous in Arikapú (rê ‘egg’, rê ‘to dance’). Finally, both ‘meat, flesh’ and ‘thorn’ are homophonous in Proto-Jabutí (*ni), Northern Jê (Apinajé ni), and Rikbaktsá (-ni).45

<table>
<thead>
<tr>
<th>JÊ FAMILY</th>
<th>JABUTÍ FAMILY</th>
<th>MEANING</th>
</tr>
</thead>
<tbody>
<tr>
<td>(50a) *ma</td>
<td>*ma</td>
<td>‘hear, know’</td>
</tr>
<tr>
<td>(50b) *ma</td>
<td>*ma</td>
<td>‘liver’</td>
</tr>
<tr>
<td>(50c) Api tf-wa</td>
<td>*tfo</td>
<td>‘tooth’</td>
</tr>
<tr>
<td>(50d) Api tf-wa</td>
<td>*tfo</td>
<td>‘to bathe’</td>
</tr>
<tr>
<td>(50e) *yr</td>
<td>ARI rê</td>
<td>‘egg’</td>
</tr>
<tr>
<td>(50f) *yr</td>
<td>ARI rê</td>
<td>‘to dance’</td>
</tr>
<tr>
<td>(50g) Api ni</td>
<td>*ni</td>
<td>‘flesh’</td>
</tr>
<tr>
<td>(50h) *j-ǐ</td>
<td>*ni</td>
<td>‘thorn’</td>
</tr>
</tbody>
</table>

45 Another example is a “minimal pair” for the contrast between *t and *r (in Proto-Jabutí and, maybe, Proto-Jê). In both Proto-Jabutí and Xavante (Central Jê), the words for ‘rope’ and ‘to sleep’ are nearly homophonous, except for the consonant in the second syllable (cf. Xavante n-ôô ‘to sleep’, n-ôô ‘rope’; Proto-Jabutí *nûôô ‘to sleep’ [see Swadesh’s no. 60], *nûrô [ARI nûrô, Dje nûrô] ‘rope made of tucuma’).
(50e) and (50f) have additional explanatory value. As mentioned in the Arikapú descriptive section above, the nasal vowel /e/ is very rare, and most of its occurrences can be explained as a consequence of nasal spreading from a following syllable, resulting in the nasalization of /e/. The only minimal pair available involves the homophonous stems re ‘egg’/ to dance’ and the classifying stem re ‘worm’. But, thanks to comparative evidence, the (apparently exceptional) nasality in this case can be explained as a leftover from a formerly existing *ŋ.

4.2. Grammatical evidence. The series of singular personal prefixes that has been reconstructed for Proto-Jabutí (table 20) is strikingly similar to the prefix series found in Jê and other well-established Macro-Jê languages, such as Maxakalí (table 21). The comparative corpora include a few examples of inflected words presenting both prefix and root cognates between both families (51); note that, in (51), the roots can be reconstructed for at least one of the proto-languages:

<table>
<thead>
<tr>
<th>PARKATEJÉ</th>
<th>ARIKAPÚ</th>
<th>MEANING</th>
</tr>
</thead>
<tbody>
<tr>
<td>intium</td>
<td>itfu</td>
<td>‘my father’</td>
</tr>
<tr>
<td>intfe</td>
<td>itfi</td>
<td>‘my mother’</td>
</tr>
<tr>
<td>ikra</td>
<td>ikraj</td>
<td>‘my child’</td>
</tr>
</tbody>
</table>

Notice that in Arikapú there is homophony between the first- and third-person prefixes; not surprisingly, such a homophony leads to ambiguity
between a first-person and a third-person “generic” reading (e.g., *i-tfawa ‘my flower’ ~ ‘(its) flower’; *i-tfapa ‘my arm’ ~ ‘(its) branch’). Such ambiguity does not occur in Djeoromitxí, where a zero-form emerged representing the first person.

Remarkably, a similar state of affairs can also be found in Jê languages. In Proto-Jê (as in Maxakalí), the first- and third-person prefixes differ phonologically only by the occurrence, with the former, of a final consonant (as illustrated by languages such as Apinajé). In languages such as Xerente (Central Jê) and Parkatêjê (Northern Jê), phonological processes leading to homophony between the first- and third-person prefixes tend to cause ambiguity and, consequently, innovations in the prefix system. Thus, in Parkatêjê, the first-person prefix is *i-, which would be homophonous with the inherited third-person marker (table 21); the adoption of a zero-prefix for the third person prevents ambiguities. The fact that Arikapú is more conservative than Djeoromitxí, preserving both prefixes despite their homophony, makes such similarities between both families rather evident.

An additional example of likely grammatical affinity between both families (the existence of “linking prefixes”) is discussed below (4.3).

4.3. Arikapú and Djeoromitxí in a comparative perspective: an evaluation. Despite the fact that Jabutí is a very small family and that both of its members are still rather similar, it is important to emphasize that Djeoromitxí and Arikapú each contribute essential bits of information to the reconstruction of Proto-Jabutí—and, consequently, to the detection of evidence of wider genetic relationships. For instance, Arikapú is phonologically more conservative, preserving, among other things: (a) vowel distinctions which were apparently lost in Djeoromitxí (*i vs. *i); (b) consonant clusters (*mr, *pr, *kr), simplified in Djeoromitxí; and (c) an independent reflex of Proto-Jabutí *tf (which, in Djeoromitxí, merged with reflexes of Proto-Jabutí *r).

The importance, for comparative studies, of such conservativeness of Arikapú, a previously undescribed language with only two speakers, cannot be stressed enough (see Moore [forthcoming]). Considering the isolating, monosyllabic nature of Jabutí—and Macro-Jê—words, one can see that any processes that lead to further loss of phonological material pose additional challenges for the comparative linguist. The fact that Arikapú preserves consonantal clusters allows the identification of correspondences which would otherwise be harder to detect (note Arikapú kraj ‘offspring’ :: Proto-Jê *kra, etc.).

In one aspect, Arikapú is also morphologically more conservative, preserving the complete series of singular personal prefixes, *i- ‘first person’, *a- ‘second person’, and *i- ‘third person’, in spite of the ambiguity resulting from the homophony between the first- and third-person markers. This makes
it more evident that the complete series of singular personal prefixes in Proto-Jabutí and Proto-Jê were cognates, which provides a stronger piece of evidence than the existence of individual personal prefixes alone. The fact that in both Arikapú and several Jê languages the first- and third-person prefixes became homophonous only adds to the strength of the proof.

There is at least one aspect in which Djeoromitxí seems to be morphologically more conservative than Arikapú. As we have seen, a number of noun and verb stems in Djeoromitxí present a word-initial morphophonemic alternation between r- ~ n- when there is a morphologically contiguous determiner, and h- elsewhere. This is a fully productive process in Djeoromitxí but not in Arikapú. However, the existence of a few, apparently frozen, relic alternations in Arikapú (e.g., haroko ~ tʃaroʃko ‘to speak’) seems to suggest that such a process can be reconstructed for Proto-Jabutí.

For those acquainted with (Macro-)Jê languages, the alternations found in Djeoromitxí are strikingly similar to the ones involving the so-called relational prefixes, a morphological peculiarity described for Jê and other Macro-Jê families (Rodrigues 2001 and Ribeiro 2004a). As the phonological correspondences show, the consonants involved in the alternations found in Djeoromitxí, which trace back to Proto-Jabutí *n- and *tʃ- (53), correspond regularly to Proto-Jê *j (52), which happens to be the consonant of the “relational prefix” in the languages of this family.46

(52a) PJê *j- ‘meat’ > Kaingáng nī, Apinajé j-ĩ, etc.
(52b) PJê *j-ua ‘tooth’ > Kaingáng ja, Apinajé tʃ-ua, etc.
(52c) PJê *j-um ‘father’ > Kaingáng jɔŋ, Parkatêjê tʃ-um, etc.
(52d) PJê *j-arkua ‘mouth’ > Kaingáng jǎŋká ‘door’, Apinajé j-akwa, etc.

(53a) PJab *nĩ ‘meat’ > Djeoromitxí nĩ, Arikapú nĩ
(53b) PJab *tʃo ‘tooth’ > Djeoromitxí rʊ ~ hʊ, Arikapú tʃukrīhã
(53c) PJab *tiʃ ‘father’ > Djeoromitxí ru ~ hu, Arikapú tiʃ
(53d) PJab *tʃako ‘mouth’ > Djeoromitxí rɑkʊ ~ hɑkʊ, Arikapú tʃako

It is likely that, in Proto-Jê, *j already had two allophones, pronounced as *[n] before nasal vowels and *[j] elsewhere. In Kaingáng, *[n] became /n/, thus merging with reflexes of Proto-Jê *n, whereas *j remains as /ʃ/. Similar splits are found throughout the family, as well as in other languages of the

46 *j also corresponds regularly to the consonant of the relational prefix in Karajá (Ribeiro 2004a) and other families. Notice that, unlike its cognates in Northern Jê languages (for instance, Parkatêjê j-ĩ ‘meat’, h-ĩ ‘its meat’), the word for ‘meat’, nĩ, does not display consonant alternation in Djeoromitxí, although such alternations are found with other stems in the same phonological environment (cf. hihōnɔnɔdə ~ nihōnɔnɔdə ‘to work’, in Pires 1992:106;124). However, the initial consonant is also invariable in Central Jê (Xavante jĩ) and Karajá (dc). It is therefore possible that Northern Jê innovated, reanalyzing the initial consonant as a prefix.
stock, such as Ofayé (cf. f-er ‘mouth’, j-ôra [nôra] ‘tongue’; cf. respectively Proto-Jê *j-ar- and *j-ôts). As we have seen, Proto-Jabutí *n and *tS (in those examples involving alternation) correspond regularly to Proto-Jê *j. It is interesting to point out that Arikapú /tS/ rarely occurs before nasal vowels. This suggests that, at a certain stage of Pre-Proto-Jabutí, *tS and *n were in complementary distribution, a situation reminiscent of the scenario reconstructed for Proto-Jê *j. Again, it is the interplay between data from Arikapú and Djeoromitxí which contributes to provide a clearer picture of Proto-Jabutí (and beyond) and of the correspondences between Jabutí and the other Macro-Jê families.

4.4. Jabutí as Macro-Jê: an appraisal. As with any claim of remote relationship, one has to be careful in order to avoid the possibility of chance similarities being misinterpreted as proof of genetic connection. In Macro-Jê, one of the most controversial language groupings in South America, this possibility is particularly hard to avoid, on account of the isolating (and mostly monosyllabic) nature of the morphemes (Rodrigues 1999; 2001, Campbell 1988:600, and Meillet 1967:55). In addition, some of the grammatical morphemes generally pointed out as evidence of genetic relationship in Macro-Jê are also found elsewhere (the so-called pan-americanisms [Campbell and Kaufman 1983:366]). Taken as a whole, however, the evidence presented here seems to strongly point to a genetic relationship. A number of regular phonological correspondences were detected, in a part of the lexicon that is generally considered as being diachronically stable and less prone to borrowing (many of the compared forms are found in Swadesh’s basic list). One major advantage in relation to other works proposing the membership of individual families in the Macro-Jê stock (Davis 1968, Boswood 1973, and Gudschinsky 1971) is that, in the present case, two proto-languages can be compared, thus reducing the possibility of including in the comparison chance similarities and loans, besides increasing the temporal depth by at least two millennia. In proposing the inclusion of Karajá and Maxakalí, Davis (1968) relies on the diagnostic value of the Swadesh list: “Maxakalí and Karajá are included in the same stock with the Jê languages on the basis of the fact that regular sound correspondences are detectable in a relatively small corpus of data and on the basis of lexical similarity. Lexicostatistical comparisons based on the Swadesh 100–word list show about 25% shared cognates between Maxakalí or Karajá and individual Jê languages.”

In the present case, even if we adopt a more conservative approach, computing only the first two categories of cognates (those which can be re-

47 Notice that Ribeiro (forthcoming) reconstructs two forms with the meaning ‘mouth’: *j-ar- and *j-arkua. Ofayé f-er is a cognate of the former.
constructed for at least one of the proto-languages), we would have a rate of approximately 25% cognation (23 out of 97, considering that pairs such as ‘hear/know’, ‘skin/bark’, ‘wood/horn’, which have different entries in Swadesh’s list, are one and the same word both in Jabutí and Jê). Therefore, based on the same kind of similarities pointed out by Davis in advocating the inclusion of Maxakalí and Karajá in the Macro-Jê stock (improved, in our case, by the careful reconstruction of the proto-languages and by the identification of grammatical similarities), it seems quite reasonable to propose the inclusion of Jabutí in the stock as well.

5. Implications for prehistory. The thesis presented in this article, if confirmed by further scrutiny, would have far-reaching consequences for theories on the prehistoric dispersion of the Macro-Jê-speaking tribes. Current scholarship on Macro-Jê rests on the assumption of a supposedly higher diversity in eastern Brazil, where the majority of the members of the stock would be located. As Urban (1998:91) points out, if families such as Kamakã, Puri, Krenâk, and Maxakalí “are only remotely related, this would be an area of great linguistic diversity . . . , and, thus, a possible place for the dispersion that took place 5 or 6 thousand years ago.” If, however, the aforementioned eastern families are closely related (as proposed by Ribeiro 2007) and additional members of the stock are proven to exist in the far west, current theories will have to be reconsidered. Central Brazil would then be a stronger candidate as the original homeland of Macro-Jê.

6. Conclusion. Although the correspondences shown above may require further refinement, they make a fairly strong case for considering the inclusion of Jabutí in the Macro-Jê stock. As the discussion in 4 suggests, there is a considerable degree of regular, recurrent correspondences, including a number of grammatical elements and a possible case of shared aberrancy. Additional research, including an investigation of the apparent cases of shared idiosyncrasies in the personal pronominal prefix systems, may reveal further regularities. Thus, the Jabutí/Macro-Jê hypothesis, first raised by Curt Nimuendajú, is confirmed by our initial findings. Further investigation of the Jabutí languages will no doubt come to enrich the field of Macro-Jê studies in the future.

48 The fact that one and the same word combines the meanings ‘to hear’ and ‘to know’ may be interesting in itself. Although the convergence between the two meanings can occur independently in unrelated languages, we should note that the use of ‘hearing’ as a metaphorical locus for ‘understanding, awareness, knowledge’ is a rather common characteristic in Macro-Jê. In Karajá, for example, ‘to be stupid’ is ‘to not have ears’; ‘to hear’ (constructed with the noun stem for ‘ear’) equals ‘to understand’; ‘to faint’ is ‘to have (one’s) ears clogged’.
## APPENDIX A

**Swadesh Basic Vocabulary for the Jabutí Languages with Proto-Jabutí and Proto-Jê Reconstructions**

Below is Swadesh’s basic list of 100 words, as represented in Bynon (1983:268). It contains Arikapú and Djeoromitxí forms, as well as reconstructed Proto-Jabutí and Proto-Jê forms wherever available.

<table>
<thead>
<tr>
<th>English</th>
<th>Arikapú</th>
<th>Djeoromitxí</th>
<th>Proto-Jabutí</th>
<th>Proto-Jê</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. I</td>
<td>ihc, i-</td>
<td>hu, Ø</td>
<td>*i-</td>
<td>*i-</td>
</tr>
<tr>
<td>2. you (sg/pl)</td>
<td>ahe, a-</td>
<td>adze, a-</td>
<td>*a</td>
<td>*a-</td>
</tr>
<tr>
<td>3. we</td>
<td>têhc, tê-</td>
<td>hiru, hi-</td>
<td>*hi</td>
<td></td>
</tr>
<tr>
<td>4. this</td>
<td>ájha</td>
<td>we</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. that</td>
<td>mwehã, nihã</td>
<td>mupc, mc, nudžu</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. who</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. what</td>
<td>ãwhã</td>
<td>hatime</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. not</td>
<td>màj</td>
<td>mâ</td>
<td>*mâj</td>
<td></td>
</tr>
<tr>
<td>8a. not, no</td>
<td>tõ</td>
<td>tô</td>
<td>*tò</td>
<td>*tô</td>
</tr>
<tr>
<td>9. all (finished)</td>
<td>hötã</td>
<td>hätzã</td>
<td>(*..tã)</td>
<td></td>
</tr>
<tr>
<td>10. many</td>
<td>hzumrã</td>
<td>hõta</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. one</td>
<td>tôjwê</td>
<td>utị</td>
<td>*pit</td>
<td></td>
</tr>
<tr>
<td>12. two</td>
<td>hêri</td>
<td>džamu</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13. big</td>
<td>títifi, rukře</td>
<td>títifi</td>
<td>*títifi</td>
<td></td>
</tr>
<tr>
<td>14. long</td>
<td>rzhóñi</td>
<td>kurtifì</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15. small</td>
<td>mroj</td>
<td>tö</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16. woman</td>
<td>pâuc</td>
<td>paku</td>
<td>*paku</td>
<td></td>
</tr>
<tr>
<td>16a. female</td>
<td>tñi</td>
<td>dżi</td>
<td>*dżi</td>
<td></td>
</tr>
<tr>
<td>17. man</td>
<td>onõhc</td>
<td>tñu</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17a. male</td>
<td>tñi, tññi</td>
<td>-ru, hutifi</td>
<td>*tñu, tññi</td>
<td>*j-um</td>
</tr>
<tr>
<td>18. person</td>
<td></td>
<td>hikomu</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19. fish</td>
<td>minù</td>
<td>minö</td>
<td>*minâ</td>
<td></td>
</tr>
<tr>
<td>20. bird</td>
<td>(āmitifì)</td>
<td>mititifù</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21. dog</td>
<td>kura</td>
<td>wa</td>
<td></td>
<td></td>
</tr>
<tr>
<td>22. louse</td>
<td>tao</td>
<td>tõdze</td>
<td></td>
<td></td>
</tr>
<tr>
<td>23. tree, wood</td>
<td>ku</td>
<td>ku</td>
<td>*ku</td>
<td>*ko</td>
</tr>
<tr>
<td>24. seed</td>
<td>hâ</td>
<td>hö</td>
<td>*hö</td>
<td>*si</td>
</tr>
<tr>
<td>25. leaf</td>
<td>-nì, tñaro</td>
<td>nì</td>
<td>*-nì</td>
<td></td>
</tr>
<tr>
<td>26. root</td>
<td>niri</td>
<td>rari</td>
<td></td>
<td>*j-are</td>
</tr>
<tr>
<td>27. bark</td>
<td>kò</td>
<td>kò</td>
<td>*kò</td>
<td>*kì</td>
</tr>
<tr>
<td>28. skin</td>
<td>kò</td>
<td>kò</td>
<td>*kò</td>
<td>*kì</td>
</tr>
<tr>
<td>29. flesh</td>
<td>nî</td>
<td>nî</td>
<td>*nî</td>
<td>*j-î</td>
</tr>
<tr>
<td>30. blood</td>
<td>tñò</td>
<td>kñi</td>
<td></td>
<td></td>
</tr>
<tr>
<td>31. bone</td>
<td>tñi, i</td>
<td>dži</td>
<td>*dži (or *i)</td>
<td>*si</td>
</tr>
<tr>
<td>32. grease</td>
<td>tuka</td>
<td>tò</td>
<td>*tù(ka)</td>
<td></td>
</tr>
<tr>
<td>33. egg</td>
<td>-rè</td>
<td>dže</td>
<td>*(. rè)</td>
<td>*qre</td>
</tr>
<tr>
<td>34. horn</td>
<td>nîpkojkù</td>
<td>imku</td>
<td>*ko</td>
<td></td>
</tr>
<tr>
<td>35. tail</td>
<td>nîñoj</td>
<td>u</td>
<td></td>
<td></td>
</tr>
<tr>
<td>36. feather</td>
<td>to</td>
<td>rari</td>
<td></td>
<td>*j-ar</td>
</tr>
<tr>
<td>37. hair</td>
<td>kai</td>
<td>kuahi, hi</td>
<td></td>
<td></td>
</tr>
<tr>
<td>38. head</td>
<td>kaj</td>
<td>kuaka</td>
<td></td>
<td>*krä</td>
</tr>
<tr>
<td>No.</td>
<td>English</td>
<td>Djeoromitxí</td>
<td>Arikapú</td>
<td>Proto-Jabutí</td>
</tr>
<tr>
<td>-----</td>
<td>---------</td>
<td>-------------</td>
<td>---------</td>
<td>--------------</td>
</tr>
<tr>
<td>39.</td>
<td>ear</td>
<td>nipwaro</td>
<td>*nipi</td>
<td>(*niπi)*⁴⁹</td>
</tr>
<tr>
<td>40.</td>
<td>eye</td>
<td>hākare⁵⁰</td>
<td>hōka</td>
<td><em>hākare</em></td>
</tr>
<tr>
<td>41.</td>
<td>nose</td>
<td>ninka</td>
<td>nikate</td>
<td>*ji-ja</td>
</tr>
<tr>
<td>42.</td>
<td>mouth</td>
<td>tfako (nuku)</td>
<td>-raku/haku</td>
<td>*tfako</td>
</tr>
<tr>
<td>43.</td>
<td>tooth</td>
<td>tjkorihā</td>
<td>(-)ru</td>
<td>*tro</td>
</tr>
<tr>
<td>44.</td>
<td>tongue</td>
<td>nuktare</td>
<td>nūt(ře)</td>
<td>*ji-ōt</td>
</tr>
<tr>
<td>45.</td>
<td>claw (i.e., nail)</td>
<td>nikataj</td>
<td>nikate</td>
<td>*nikataj</td>
</tr>
<tr>
<td>46.</td>
<td>foot</td>
<td>praj</td>
<td>pa, panikə</td>
<td>*praj</td>
</tr>
<tr>
<td>47.</td>
<td>knee</td>
<td>mēpe</td>
<td>pepe</td>
<td>*mēpe⁵¹</td>
</tr>
<tr>
<td>48.</td>
<td>hand</td>
<td>nīkaj, nīku, niku</td>
<td>nīhu, nihu</td>
<td>*nī, u⁵²</td>
</tr>
<tr>
<td>49.</td>
<td>belly</td>
<td>prika</td>
<td>pika</td>
<td>*prika</td>
</tr>
<tr>
<td>50.</td>
<td>neck</td>
<td>poko, (kopo)⁵³</td>
<td>kōrih</td>
<td>*poko</td>
</tr>
<tr>
<td>50a.</td>
<td>neck2</td>
<td>ruwaj</td>
<td>ru</td>
<td></td>
</tr>
<tr>
<td>51.</td>
<td>breasts</td>
<td>nuni</td>
<td>nowi</td>
<td>*nuni</td>
</tr>
<tr>
<td>52.</td>
<td>heart</td>
<td>moka</td>
<td>matutuka</td>
<td>*m(tutuka)</td>
</tr>
<tr>
<td>53.</td>
<td>liver</td>
<td>mə̆</td>
<td>mə̆</td>
<td>*ma</td>
</tr>
<tr>
<td>54.</td>
<td>drink</td>
<td>o</td>
<td>nō</td>
<td></td>
</tr>
<tr>
<td>55.</td>
<td>eat</td>
<td>pu</td>
<td>pu</td>
<td>*pu</td>
</tr>
<tr>
<td>55a.</td>
<td>eat2</td>
<td>ku</td>
<td>ku</td>
<td>*ku</td>
</tr>
<tr>
<td>56.</td>
<td>bite</td>
<td>ku</td>
<td>ku</td>
<td>*ku</td>
</tr>
<tr>
<td>57.</td>
<td>see</td>
<td>arā</td>
<td>mō̆</td>
<td>*arā</td>
</tr>
<tr>
<td>58.</td>
<td>hear</td>
<td>mə̆</td>
<td>mə̆</td>
<td>*ma</td>
</tr>
<tr>
<td>59.</td>
<td>know</td>
<td>mə̆</td>
<td>mə̆</td>
<td>*ma</td>
</tr>
<tr>
<td>60.</td>
<td>sleep</td>
<td>nūtō</td>
<td>nōtō</td>
<td>*nūtō</td>
</tr>
<tr>
<td>61.</td>
<td>die</td>
<td>pi</td>
<td>hahi</td>
<td></td>
</tr>
<tr>
<td>62.</td>
<td>kill (shoot)</td>
<td>konə</td>
<td>hi, hiroku</td>
<td></td>
</tr>
<tr>
<td>62a.</td>
<td>kill (club)</td>
<td>tōmrə̆</td>
<td>tumi</td>
<td>*t.mr.</td>
</tr>
<tr>
<td>63.</td>
<td>swim</td>
<td>(i)pri</td>
<td>iwa</td>
<td></td>
</tr>
<tr>
<td>64.</td>
<td>fly</td>
<td>tfo po</td>
<td>irariku</td>
<td></td>
</tr>
<tr>
<td>65.</td>
<td>walk</td>
<td>kōraj</td>
<td>dżikre</td>
<td>(*kōraj)*⁵⁴</td>
</tr>
<tr>
<td>66.</td>
<td>come (arrive)</td>
<td>prə̆</td>
<td>pe</td>
<td>*prə̆</td>
</tr>
<tr>
<td>67.</td>
<td>lie</td>
<td>kōrāj</td>
<td>ure</td>
<td>(*..rə̆)</td>
</tr>
<tr>
<td>68.</td>
<td>sit</td>
<td>nō</td>
<td>hunə</td>
<td>*j-ō</td>
</tr>
<tr>
<td>69.</td>
<td>stand</td>
<td>tʃuc</td>
<td>tumə̆</td>
<td></td>
</tr>
<tr>
<td>70.</td>
<td>give</td>
<td>ū</td>
<td>ū</td>
<td>*ū</td>
</tr>
<tr>
<td>71.</td>
<td>say</td>
<td>tʃaroko</td>
<td>piro, rumə</td>
<td></td>
</tr>
<tr>
<td>72.</td>
<td>sun</td>
<td>tɔhə̆</td>
<td>tōhō</td>
<td>*tōhə̆</td>
</tr>
</tbody>
</table>

⁴⁹ The Arikapú form is probably a formerly productive combination that included *niπi ‘ear’.
⁵⁰ The words for ‘knee’ seem partially cognate. We may assume that original word-initial *m merged with /p/ in Djeoromitxí under denasalization, whereas it was transferred as expected in Arikapú.
⁵¹ The resemblance to the Kwaza (isolate) verb kerai- ‘to go, leave, walk’ is striking and may be a reason not to reconstruct this item.
⁵² This alternative pronunciation is probably due to influence from Djeoromitxí, where metathesis of syllables must have occurred.
⁵³ The element -re is possibly a lexicalized suffix with classifying properties.
⁵⁴ The element -re is possibly a lexicalized suffix with classifying properties.
⁵⁵ The element -re is possibly a lexicalized suffix with classifying properties.
570 INTERNATIONAL JOURNAL OF AMERICAN LINGUISTICS

73. moon kupa kupa (*kupa)
74. star wirwire bzirobezi (*
74a. big star warwaro kurawalji
75. water, liquid mi bziromi (ru)
75a. water, liquid i i (*i)
76. rain nág nipa (*
76a. rain (V) roko hok (*roko)
77. stone kra ta *kra *ken
78. sand kikira nomaru
79. earth mĩ(ka) mĩ(ka) *mĩ(ka) *pika
80. cloud mē meks(ken) (*mē(ken)
81. smoke tžio ptieno
82. fire piko ptie (*ptie
83. ash pikompo ptiekam (*pti(e(mo)
84. burn kono tštu (*tšu(*
85. path wi wi (*wi
86. mountain kamĩ mri
87. red nũro nũ (*nũ(3)
88. green kapi kapi (*kapi
89. yellow numu buzim (*numu56
90. white māo kō(mu)
91. black kario miru
92. night patjiu patjiro (*patjiro
93. hot kō tš (*tš
94. cold (k)šiti (dziidziru)
95. full maz dzi(zi)
96. new kamu kamu (*kamu
96a. new2 nute (inan)
96b. young girl nunika noki (*nuni
97. good hāwi madž (*
98. round ka ka (*ka
99. dry karo kun (*karo
100. name tafiū tōhī (*tōhī

55 Note that Tupí languages tend to have a similar form.
56 A very speculative explanation for the similarities is found in van der Voort (2007:159). It involves omission of a final glide in both languages, as also observed in *tšamu, ARI tšamu, Dje habzi 'cotton' (n. 16), combined with syllable metathesis in Djeoromix, as also hypothesized in n. 53.
REFERENCES


BONTKES, WILLEM. 1968. [Vocabulary of Arikapú]. Ms. (10 pages), Archive, Summer Institute of Linguistics, Porto Velho.


_____ 1955b. [Vocabulary of Arikapú]. Ms. (51 pages), Zürich.


Moore, Denny, and Hein van der Voort, eds. Forthcoming. Amazonian Linguistic Stocks and Their Prehistory.


OPPORTUNITIES


_____. 2004b. Uma hipótese sobre a origem do padrão ergativo em algumas línguas Jê setentrionais. Ms. (5 pages), University of Chicago.


_____. Forthcoming. A reconstruction of Proto-Jê (and its consequences for the Macro-Jê hypothesis).


_____. Forthcoming a. Caspar was right: Mashubi is Arikapú.

_____. Forthcoming b. The Arikapú language: Grammatical sketch and dictionary.


