Provided for non-commercial research and educational use only.
Not for reproduction or distribution or commercial use

This article was originally published in the *Encyclopedia of Language & Linguistics, Second Edition*, published by Elsevier, and the attached copy is provided by Elsevier for the author's benefit and for the benefit of the author's institution, for non-commercial research and educational use including without limitation use in instruction at your institution, sending it to specific colleagues who you know, and providing a copy to your institution's administrator.

All other uses, reproduction and distribution, including without limitation commercial reprints, selling or licensing copies or access, or posting on open internet sites, your personal or institution's website or repository, are prohibited. For exceptions, permission may be sought for such use through Elsevier's permissions site at:

http://www.elsevier.com/locate/permissionusematerial

Ribeiro E R (2006), Macro-Jê. In: Keith Brown, (Editor-in-Chief) *Encyclopedia of Language & Linguistics, Second Edition*, volume 7, pp. 422-426. Oxford: Elsevier.

Weaver W (1955). 'Translation.' In *Machine translation of languages*. Cambridge, MA: MIT Press.
Yngve V (1957). 'A Framework for Syntactic Translation.' *Mechanical Translation 4(3)*, 59–65; reprinted in Nirenburg S, Somers H & Wilks Y (eds.) (2003). *Readings in machine translation*. MIT Press. 39–44.
Zarechnak M (1959). 'Three levels of linguistic analysis.' *Journal of the Association for Computing Machinery 6(1)*, 24–32.

Relevant Website

http://www.ldc.upenn.edu – Linguistic Data Consortium.

Macro-Jê

E R Ribeiro, Museu Antropológico, Universidade Federal de Goiás, Goiás, Brazil

© 2006 Elsevier Ltd. All rights reserved.

The Macro-Jê stock comprises the Jê family and a number of possibly related language families, all of which are located in Brazil. Macro-Jê is arguably one of the lesser-known language groups of South America, its very existence as a genetic unit being still "a working hypothesis" (Rodrigues, 1999: 165). According to Rodrigues (1986, 1999), whose classification is the most widely accepted among researchers working on Brazilian languages, the 'Macro-Jê hypothesis' comprises 12 different language families: Jê, Kamakã, Maxakalí, Krenák, Purí, Karirí, Yatê, Karajá, Ofayé, Boróro, Guató, and Rikbaktsa. The existence of Jê as a language family has been recognized since early classifications of South American languages (Martius, 1867). 'Jê' is a Portuguese spelling for a Northern Jê collective morpheme ([je] in Apinajé, for instance) that occurs in the names of several Jê-speaking peoples. The term 'Macro-Jê' was coined by Mason (1950), replacing earlier labels, such as 'Tapuya' and 'Tapuya-Jê.'

Comparative Evidence

Recent classifications (Rodrigues, 1986; Greenberg, 1987; Kaufman, 1994) differ as to the precise scope of Macro-Jê, although there is agreement on the inclusion of most of the families (Table 1). Except for Karirí (included only by Rodrigues), Greenberg and Kaufman included all the families listed above. In addition, Greenberg included Chiquitano (also included by Kaufman), Jabutí, and Otí. Given the lack of comprehensive comparative studies, the Macro-Jê status of some of these families is still an open question. Although Guató is included in the stock by all of the aforementioned classifications, a case for its inclusion has yet to be made, beyond the superficial, inconclusive evidence presented so far (Rodrigues, 1986, 1999). On the other hand, a preliminary comparison has revealed compelling evidence for the inclusion of the Jabutí family into the Macro-Jê

Table 1 The Macro-Jê Hypothesis[a]

1. **Jê**
   †Jeikó
   Northern Jê: Panará, Suyá, Kayapó, Timbíra (Parkatêjê, Pykobjê, etc.), Apinajé
   Central Jê: Xavánte, Xerénte, †Akroá-Mirim, †Xakriabá
   Southern Jê: Kaingáng, Xokléng, †Ingaín
2. **Kamakã**
   †Kamakã, †Mongoyó, †Menién, †Kotoxó, †Masakará
3. **Maxakalí**
   Maxakalí, †Pataxó, †Kapoxó, †Monoxó, †Makoní, †Malalí
4. **Krenák**
   Krenák (Botocudo, Borúm)
5. **Purí (Coroado)**
   †Coroado, †Purí, †Koropó
6. **Ofayé**
   Ofayé
7. **Rikbaktsá**
   Rikbaktsá
8. **Boróro**
   Boróro, †Umutína, †Otúke
9. **Karajá**
   Karajá (including four dialects, Southern Karajá, Northern Karajá, Javaé, and Xambioá)
10. **Karirí**
    †Kipeá, †Dzubukuá, †Pedra Branca, †Sabuyá
    (included by Rodrigues but not Greenberg or Kaufman)
11. **Jabutí**
    Djeoromitxí (Jabutí)
    Arikapú
    (included by Greenberg but not Rodrigues or Kaufman)
12. **Yatê**
    Yatê
13. **Guató**
    Guató
14. **Chiquitano**
    Chiquitano (Besiro)
    (included by Greenberg and Kaufman, but not Rodrigues)
15. **Otí**
    †Otí (Eo-Xavánte)
    (the inclusion of Otí, proposed only by Greenberg, is not substantiated by the available data)

[a]Extinct languages are indicated by †. Based on Greenberg, 1987; Rodrigues, 1986, 1999; Kaufman, 1994.

Author's Personal Copy

stock (Voort and Ribeiro, 2004), thus corroborating a hypothesis suggested in the 1930s by ethnographer Curt Nimuendaju (Nimuendaju, 2000: 219–221). Greenberg's main piece of evidence for the inclusion of Chiquitano was the entire set of singular personal prefixes (Greenberg, 1987: 44), which are strikingly similar to the ones found in several Macro-Jê families; convincing lexical evidence, however, has not been presented thus far. As for Otí, a poorly documented language once spoken in southern Brazil, the meager available data do not support its inclusion in the Macro-Jê stock.

The only family-level reconstruction available is Davis (1966), for Proto-Jê. 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; Davis, 1968), Purí (Loukotka, 1937), Boróro (Guérios, 1939), Krenák (Loukotka, 1955; Seki, 2002), Karajá (Davis, 1968), Ofayé (Gudschinsky, 1971), Rikbaktsá (Boswood, 1973), and Jabutí (Voort and Ribeiro, 2004). In addition, some studies have shown very suggestive cases of morphological idiosyncrasies shared by Jê, Boróro, Maxakalí, Karirí, Karajá, and Ofayé (Rodrigues, 1992, 2000b). Thus, although the inclusion of many of the families into the Macro-Jê stock is being further corroborated by additional research, for others (namely Guató, Chiquitano, and Yatê) the hypothesis has yet to be systematically tested. The precise relationship among the suggested members of the stock also remains to be worked out.

Long-Range Affiliations

Greenberg (1987) suggested that Macro-Jê would be related to his Macro-Pano and Macro-Carib stocks, as part of a Jê-Pano-Carib branch of 'Amerind.' However, as Rodrigues (2000a) pointed out, Greenberg's purported evidence does not withstand careful examination. Rodrigues (1985, 2000a) proposed instead a relationship between Tupí, Carib, and Macro-Jê, noting grammatical and lexical similarities among the three language groups (especially between Carib and Tupí). Davis (1968) also mentioned a few lexical similarities between Proto-Jê and Proto-Tupí. Although the evidence presented so far suggests that Rodrigues's proposal is more plausible than Greenberg's, any hypothesis of distant genetic relationship at such a level must be considered with caution. Considering that the precise boundaries of Macro-Jê are still uncertain, much more research at the family and stock levels needs to be conducted before such long-range classifications can be proposed on solid scientific grounds.

Location

All Macro-Jê languages are spoken in Brazilian territory, although in the past Otúke (Boróro) and Ingaín (Southern Jê), both now extinct, were spoken in Bolivia and Argentina, respectively. Chiquitano, listed as a Macro-Jê language by Greenberg (1987) and Kaufman (1994), is also spoken in Bolivia, as well as in Mato Grosso, Brazil. Although the Jabutí languages and Rikbaktsá are spoken in the southern fringes of the Amazon (Rondônia and northern Mato Grosso, respectively), 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 of Purí, Kamakã, and Karirí (all now extinct). Central and Northern Jê tribes, 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. Since several purported Macro-Jê languages were spoken in eastern Brazil, a number of them became extinct early on, under the impact of European colonization. Yatê is a remarkable exception, being the only surviving indigenous language in the Brazilian northeast.

Whereas Guató, Rikbaktsá, Karajá, Krenák, and Ofayé are all single-member families (Table 1), the Jê family has a relatively large number of members, for most of which a fair amount of descriptive material is now becoming available (mostly as graduate theses and dissertations in Brazilian universities). Ofayé has around a dozen speakers, although it is mistakenly listed as extinct by some sources (including earlier editions of Ethnologue). Boróro and Maxakalí are the only surviving languages of their respective families. All the languages of the Kamakã, Purí, and Karirí families are now extinct. While documentation on Kamakã and Purí languages consists only of brief wordlists, the Karirí languages Kipeá and Dzubukuá were documented in catechisms (Mamiani, 1698; Bernardo de Nantes, 1709; respectively) and, for Kipeá, a grammar (Mamiani, 1699) – the only published grammar of a non-Tupí language from colonial Brazil. Thus, among the extinct Macro-Jê families, Karirí is the only one for which detailed grammatical information is available. Many of the languages

Author's Personal Copy

included in the Macro-Jê stock are seriously endangered (Guató, Ofayé, Krenák, and Arikapú are especially so).

Characteristics

When compared with languages of other lowland South American families (such as Carib and Tupí-Guaraní), Macro-Jê languages typically present larger vowel inventories. For instance, Davis (1966) reconstructed, for Proto-Jê, a system of nine oral and six nasal vowels, as well as 11 consonants. Syllabic patterns are rather simple, obstruent clusters being uncommon. Stress is generally predictable. Phonologically contrastive tone oppositions occur in Yatê and Guató (Palácio, 2004). Processes such as nasal spreading and vowel harmony are generally absent. An exception is Karajá, which presents advanced tongue root vowel harmony, a rare phenomenon among South American languages (Ribeiro, 2002a). Another remarkable feature of Karajá is the existence of systematic differences between male and female speech. Female speech is more conservative, male speech being characterized, in general, by the deletion of a velar stop occurring in the corresponding female speech form (as a result of consonant deletion, vowel assimilation and fusion may also occur). This is a very productive process, applying even to loanwords (Table 2).

Most Macro-Jê languages have a relatively simple morphology. In most languages (including those of the Jabutí, 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. Tense and aspect distinctions are generally conveyed by particles and auxiliaries rather than by inflections (with few apparent exceptions, such as Yatê; cf. Costa, 2004). Noun incorporation is rare, having been reported for a few Northern Jê languages, such as Panará (which also presents postposition incorporation; cf. Dourado, 2002).

In languages with a more robust morphology, such as Karajá, Guató, and Yatê, inflectional morphology tends to be more complex with verbs than with nouns. In Karajá, for example, the verb form includes subject-agreement, voice (transitive, passive, and antipassive), and directional markers ('thither' versus 'hither'), which can be used with evidential purposes (Ribeiro, 2002b); on the other hand, the only category for which nouns inflect is possession (as in most Macro-Jê languages).

The majority of the purported Macro-Jê languages are verb final, with postpositions instead of prepositions and possessor-possessed order in genitive constructions (the exceptions being Guató, Chiquitano, and Karirí). Macro-Jê languages seemingly lack the adjective as an independent part of speech, with adjectival meanings being expressed by nouns or descriptive verbs. Oliveira (2003) offered an in-depth discussion of the properties displayed by 'descriptives' in a particular Macro-Jê language, Apinajé, illustrating well the issues involved in determining part-of-speech membership in languages in which most inflectional properties tend to be shared by nouns, verbs, and adpositions. In attributive constructions, descriptives follow the word they modify.

Languages such as Maxakalí, Karirí, and Panará are described as being predominantly ergative. In addition, a number of Jê languages are described as presenting an ergative split of some sort. That is the case of Xokléng (Urban, 1985) and Northern Jê languages such as Kayapó (Silva and Salanova, 2000) and Apinajé. Among the latter, however, ergativity seems to be rather epiphenomenal, being found only in constructions involving nominalized verbs (such as relative clauses; cf. Oliveira, 2003). Syntactic ergativity is rarely found in Macro-Jê, with the exception of Karirí, in which all grammatical criteria (verb inflection, relativization, switch-reference, word order) point to the absolutive argument (S/O) as being the syntactic pivot (Larsen, 1984).

Table 2 Female versus male speech distinctions in Karajá

Female speech	*Male speech*	
kɔwɔrʊ	*ɔwɔrʊ*	'wood'
dɪkarə̃	*dɪarə̃*	'I'
kɔhã	*ɔhã*	'armadillo'
kədə̃ra	*ədə̃ra*	'sand'
ruku	*ru*	'night'
bɛraku	*bero*	'river'
ɗəkɪ	*ɗi*	'3rd person pronoun'
kõbəra	*õbəra*	'to buy' (from Portuguese *comprar*)
kabɛ	*abɛ*	'coffee' (from Portuguese *café*)
bə̃kawa	*bə̃awa*	'firearm' (from Língua Geral *mokə̃áwa*)

Further Reading

For information on the main literature on Macro-Jê languages, including an overview of their phonological and grammatical characteristics and a short list of possible Macro-Jê cognate sets, see Rodrigues (1999). Proceedings of recent conferences (the 'Encontros Macro-Jê,' which have been taking place periodically since 2001) help to provide an updated picture of Macro-Jê scholarship; the proceedings of the first

Author's Personal Copy

two meetings were published as Santos and Pontes (2002) and D'Angelis (2004), respectively. Population figures for all Macro-Jê groups (including those now monolingual in Portuguese) can be found in Ricardo (2001).

See also: Brazil: Language Situation; Cariban Languages; Tupian Languages.

Bibliography

Bernardo de Nantes R P Fr. (1709). *Katecismo indico da lingua Kariris*. Lisbon: Valentim da Costa. [Facsimile reproduction by Platzmann J (ed.). *Catecismo da lingua Kariris*. Leipzig: Teubner, 1896.]

Boswood J (1973). 'Evidências para a inclusão do Aripaktsá no filo Macro-Jê.' In Bridgeman L I (ed.) *Série Lingüística*, 1. Brasília: SIL. 67–78.

Costa J da (2004). *'Morfologia do verbo em Yaathe.'* In D'Angelis (org.) 149–161.

D'Angelis W da R (org.) (2004). *LIAMES (Línguas Indígenas Americanas) 4. Special Issue, Proceedings of the 2nd 'Encontro de Pesquisadores de Línguas Jê e Macro-Jê,' Campinas, May 2002*.

Davis I (1966). 'Comparative Jê phonology.' *Estudos Lingüísticos: Revista Brasileira de Lingüística Teórica e Aplicada 1*, 2.10–24.

Davis I (1968). 'Some Macro-Jê relationships.' *International Journal of American Linguistics 34*, 42–47.

Dourado L (2002). 'Construções aplicativas em Panará.' *DELTA 18*, 203–231.

Greenberg J H (1987). *Languages in the Americas*. Stanford: Stanford University Press.

Gudschinsky S C (1971). 'Ofayé-Xavante, a Jê language.' In Gudschinsky S C (ed.) *Estudos sôbre línguas e culturas indígenas*. Brasília: SIL. 1–16.

Guérios R F M (1939). 'O nexo lingüístico Bororo-Merrime-Caiapó.' *Revista do Círculo de Estudos Bandeirantes 2*, 61–74.

Kaufman T (1994). 'The native languages of South America.' In Asher R E & Moseley C (eds.) *Atlas of the world's languages*. London: Routledge. 46–76; 14–25 (maps).

Larsen T W (1984). 'Case marking and subjecthood in Kipeá, Kirirí.' *Berkeley Linguistic Society Proceedings 10*, 189–205.

Loukotka C (1931). 'La familia lingüística Mašakali.' *Revista del Instituto de Etnología de la Universidad Nacional de Tucumán 2*, 21–47.

Loukotka C (1932). 'La familia lingüística Kamakan del Brasil.' *Revista del Instituto de Etnología de la Universidad Nacional de Tucumán 2*, 493–524.

Loukotka C (1937). 'La familia lingüística Coroado.' *Journal de la Société des Américanistes de Paris* n.s. *29, 157–214*.

Loukotka C (1939). 'A língua dos Patachos.' *Revista do Arquivo Municipal de São Paulo 55*, 5–15.

Loukotka C (1955). 'Les indiens Botocudo et leur langue.' *Lingua Posnaniensis 5*, 112–135.

Mamiani L V (1698). *Catecismo da doutrina Christãa na lingua brasilica da nação Kiriri*. Lisbon: Miguel Deslandes. [Facsimile:. Garcia R (ed.). *Catecismo Kiriri*. Rio de Janeiro: Imprensa Nacional, 1942.]

Mamiani L V (1699). *Arte de grammatica da lingua brasilica da naçam Kiriri*. Lisbon: Miguel Deslandes. [2nd edn.: Rio de Janeiro: Bibliotheca Nacional, 1877.]

Martius K F P von (1867). *Beiträge zur Ethnographie und Sprachenkunde Amerikas zumal Brasiliens* (2 vols). Leipzig: Friedrich Fleischer.

Mason J A (1950). 'The languages of South American Indians.' In Steward J H (ed.) *Handbook of South American Indians*, vol. 6. Washington, DC: Smithsonian Institution. 157–317.

Nimuendaju C (2000). *Cartas do Sertão, de Curt Nimuendajú para Carlos Estevão de Oliveira*. Lisbon: Museu Nacional de Etnologia/Assírio & Alvim.

Oliveira C C de (2003). 'Lexical categories and the status of descriptives in Apinajé.' *International Journal of American Linguistics 69*, 243–274.

Palácio A (2004). *'Alguns aspectos da língua Guató.'* In D'Angelis (org.). 163–170.

Ribeiro E R (2002a). 'Directionality in vowel harmony: the case of Karajá.' *Berkeley Linguistics Society Proceedings 28*, 475–485.

Ribeiro E R (2002b). 'Direction in Karajá.' In Fernández Z E & Ciscomani R M O (eds.) *Memorias, VI Encuentro de Lingüística en el Noroeste*, vol. 3. Hermosillo: Editorial UniSon. 39–58.

Ricardo C A (ed.) (2001). *Povos indígenas no Brasil: 1996/2000*. São Paulo: Instituto Socioambiental.

Rodrigues A (1985). 'Evidence for Tupí-Carib relationships.' In Klein H E & Stark L R (eds.) *South American Indian languages: retrospect and prospect*. Austin: University of Texas Press. 371–404.

Rodrigues A (1986). *Línguas brasileiras: para o conhecimento das línguas indígenas*. São Paulo: Edições Loyola.

Rodrigues A (1992). 'Um marcador Macro-Jê de posse alienável.' In *Anais da 44a. Reunião Anual da SBPC*. São Paulo: SBPC. 386.

Rodrigues A (1999). 'Macro-Jê.' In Dixon R M W & Aikhenvald A (eds.) *The Amazonian languages*. Cambridge: Cambridge University Press. 164–206.

Rodrigues A (2000a). 'Ge-Pano-Carib versus Jê-Tupí-Karib: sobre relaciones lingüísticas prehistóricas en Sudamérica.' In Miranda L (ed.) *Actas del I Congreso de Lenguas Indígenas de Sudamérica* (tomo I). Lima: Universidad Ricardo Palma. 95–104.

Rodrigues A (2000b). 'Flexão relacional no tronco Macro-Jê.' *Boletim da Associação Brasileira de Lingüística 25*, 219–231.

Santos L dos & Pontes I (eds.) (2002). *Línguas Jê: estudos vários*. Londrina: Editora UEL.

Seki L (2002). 'O Krenak (Botocudo/Borum) e as línguas Jê.' In Santos & Pontes (eds.). 15–40.

Silva M A R & Salanova A (2000). Verbo y ergatividad escindida en Mebêngôkre. In Voort H van der & Kerke S

Author's Personal Copy

van der (eds.) *Indigenous Languages of Lowland South America*. Leiden: University of Leiden. 225–242.
Urban G (1985). 'Ergativity and accusativity in Shokleng (Gê).' *International Journal of American Linguistics 51*, 164–187.
Voort H van der & Ribeiro E R (2004). 'The westernmost branch of Macro-Jê.' Paper presented at the 2nd Workshop 'Linguistic Prehistory in South America,' September 4–10. University of Oregon, Department of Linguistics.

Macrostructure

M M Louwerse and A C Graesser, University of Memphis, Memphis, TN, USA

© 2006 Elsevier Ltd. All rights reserved.

Macrostructures are structures that organize texts globally, just as microstructures organize them locally. Given that texts are not just concatenations of sentences, texts need to be structured both locally (connections between clauses and sentences) and globally (larger fragments of discourse, e.g., paragraphs). Syntactic rules, the meaning of the words of the sentence, and general heuristics of discourse form the microstructure of a text. These microstructures organize sets of interrelated propositional representations of the phrases, clauses, and sentences of the text. Macrorules translate these sequences of propositions into a smaller set of more general propositions by deleting propositions that are less important for the overall meaning of the text, by generalizing propositions into supersets and by constructing new text units that replace the meaning of the old set. Macrorules operate recursively, so that macrostructures that are formed by macrorules may be subject to another cycle of macrorules, further generalizing the gist of the text. Macrostructures are therefore abstract semantic descriptions of the semantic content of the text, similar to the text's global meaning and theme and providing global coherence.

The term 'macrostructure' for global principles of text organization was first proposed by Bierwisch in 1965 for narrative structures in literary texts. In 1968, Harris discussed a similar idea of global text structuring. In that same year the *Morphology of the folktale* by Vladimir Propp was translated from Russian (original 1928). Propp argued that Russian fairy tales share a particular narrative structure. Around that same time, narratologists such as Greimas (1966), Bremond (1964), Labov and Waletzky (1967), Lévi-Strauss (1960), and Todorov (1971) proposed similar narrative grammars. According to these grammars, stories are like sentences in that their narrative structures are structures analogous to syntactic structures. The development of the concept of macrostructures should be seen against the background of the developments of narrative structures.

'Macrostructures' became an established term in text linguistics after being further developed by Van Dijk (1972). Van Dijk's text linguistic approach was very much based on theoretical linguistics and Chomsky's (1957, 1965) generative-transformational grammar. Chomsky (a student of Harris) argued that sentences have a recursive capacity. Each sentence has a deep structure that is interpreted by the semantic component of the grammar. Syntactic transformations relate the deep structure to the surface structure of the sentence. Some of the elementary transformations consist of adjoining, moving, deleting, and copying constituents. Van Dijk argued that text grammars also have these deep and surface structures. The equivalent to the sentential surface structures are microstructures; the equivalent to the sentential deep structures are macrostructures. As with the sentential surface structures, microstructures have underlying rules to represent the underlying semantic representation of the sentences. As with the sentential deep structures, macrostructures have an abstract semantic character and are specified by macrosemantic rules operating on the microstructures. Although the concept of macrostructures remained the same, the direct link to sentence structures faded in later work (Van Dijk, 1977, 1980).

Around the time of Van Dijk's introduction of text grammars in (text) linguistics and poetics, Kintsch (1974) argued that cognitive psychology should not focus on isolated sentences only, but must focus on texts. Kintsch proposed that the representation of texts in memory is a network of interrelated propositions. These propositions are units of meaning roughly corresponding to phrases or clauses. A proposition consists of a predicate that modifies one or more arguments. A sentence like *The teacher explained the concept to the students* can be represented propositionally as (EXPLAIN, TEACHER, CONCEPT, STUDENT). It contains one predicate (EXPLAIN) and three arguments (TEACHER, CONCEPT, STUDENT). Arguments are generally the nouns of a clause, but can also be prepositional phrases and

Author's Personal Copy