

SOME MACRO-JÊ RELATIONSHIPS

IRVINE DAVIS

SUMMER INSTITUTE OF LINGUISTICS

0. Introduction
1. The Jê-Maxakalí-Karajá relationship
2. Composition of the Macro-Jê stock

0. A number of closely related languages spoken in central Brazil have been traditionally regarded as constituting the Jê linguistic family.¹ More recently the Kaingang dialects of southern Brazil, including Xokleng, have been included in this family. In a previous paper² I examined the comparative phonology of the Jê languages and attempted some reconstructions of Proto Jê lexical items. In the present paper I present evidence for the relationship of the Jê languages to two other Brazilian languages, Maxakalí and Karajá, and make some observations regarding the composition of the Macro-Jê stock.

1. Maxakalí has generally been included within the Macro-Jê stock by comparativists, but its relationship to the Jê languages has not previously been demonstrated by careful comparative techniques. Karajá, on the other hand, has generally been placed outside of Macro-Jê in spite of the fact that significant similarities to certain Jê languages were noted as long ago as by nineteenth century ethnologists.³ The paucity of

¹ The spelling of language and tribal names in this paper closely follows current Brazilian usage. The symbol x indicates a voiceless alveopalatal sibilant and j represents its voiced counterpart. Other consonantal symbols have their normal values. Accent marks are placed over stressed vowels. In addition, these marks distinguish stressed mid open é and ó from stressed mid close ê and ô.

² Irvine Davis, *Comparative Jê Phonology*, *Estudos Lingüísticos* 1.2.10-25 (1966).

³ Paul Ehrenreich, *Materialien zur Sprachkunde Brasiliens, II Die Sprache der Caraya*

available descriptive material for Karajá has undoubtedly prevented recent authors from following this lead.

Evidence for the relationship of Maxakalí and Karajá to Jê is presented below.⁴ The cognate sets have been carefully chosen and have a relatively high degree of reliability. The possibility of chance similarities or of similarities due to borrowing is reduced by comparing Maxakalí and Karajá items with forms reconstructed from divergent branches of the Jê family rather than with items from individual Jê languages. The phonetic content of each form is at least partly explainable in terms of regular correspondences, and all proposed cognates are close semantic equivalents. Furthermore, use is made of basic vocabulary items that are less likely than culture-bound references to be affected by loans.

Items preceded by an asterisk in the following list are Proto Jê reconstructions. Maxakalí, M, and Karajá, K, cognates follow. The notation K (w) indicates that the citation is from Karajá women's speech. In general, the corresponding form in men's speech lacks the velar stop. Parentheses enclose portions of Maxakalí and Karajá forms which apparently do not enter into

(Goyaz), *Zeitschrift für Ethnologie* 26.20-37, 49-60 (1894).

⁴ The Karajá data are from David Fortune of the Summer Institute of Linguistics who collaborated with the author in lining up possible Jê-Karajá cognates. The Maxakalí data are from material filed at the Museu Nacional, Rio de Janeiro by Harold Popovich of the Summer Institute of Linguistics. The Jê items are my reconstructions based on a comparison of Canela, Apinajé, Suyá, Xavante, and Kaingang. The data for these languages were collected largely by members of the Summer Institute of Linguistics and are on file at the Museu Nacional.

the comparison. Hyphens indicate infra-word morpheme divisions that are well supported by internal evidence. English glosses following the Proto Jê items apply also to the Maxakalí and Karajá cognates unless otherwise indicated.

1. * a- *your*, M ?ã-, K ã-
2. * cwa *tooth*, M -coc, K čuu
3. * ka *you*, K kai
4. * ka-kre, -kreñ *to scratch*, K i-θe
5. * kaŋã *snake*, M kãñã(noc)
6. * ka-zo, -zor *to suck*, M -cip, K -d^oo-
7. * kə *skin, bark*, M -cac, K (w) (d^oə)ki
8. * keckwa *sky*, M ñãñko(te?)
9. * -ke, -kec *left*, M -cac
10. * ken *stone*, M cap
11. * ki *hair*, M -ce?
12. * ko *horn*, M kip
13. * -kō, -kōm *to drink*, M -co?op, K -ō-
14. * kok *wind*, K (w) kihi
15. * kōn *knee*, M -kopa(-cic), K (d^oi-)əho
16. * kra *child*, M kitok *son*
17. * krā, krāñ *head*, K ra(-d^oi)
18. * krē, krēr *to eat*, M -cit, K (-rə)θō- *to swallow*
19. * krə *rotten*, M -ktoc, K rə
20. * kriz *parrot*, M konññ
21. * ku, kur *to eat*, K (w) -ki-
22. * ku-krit *tapir*, M (?i-)citta?, K (w) kōri
23. * kūm *smoke*, M -ŋōñ
24. * ku-zi *fire*, M kicap, K (heo)d^oi
25. * kwir *manioc*, M kon, kohot, K (əji)ura
26. * ma *liver*, K baa
27. * -ma, -mar *to hear*, M (-ci)pak
28. * mec *good*, M -mac
29. * meñ *honey*, K bædi
30. * mī, mīñ *alligator*, M mǎ?ãñ
31. * mō, mōr *to go, walk*, M mōŋ
32. * mrə, mrəc *ashes*, M pitohok, K bri(bi)
33. * mut *neck*, K bed^oə *throat*
34. * na *rain*, M tehac
35. * nō, nōr *to lie*, K rōrō- *to sleep*
36. * nə *eye*, K rue
37. * niw *new*, M -tip
38. * ñī *meat*, M (coŋ)ñññ, K -de
39. * ñī-ña-kre *nose*, K deãθō
40. * ñōt *to sleep*, M -ñōn
41. * ñō-tə *tongue*, M -ñōñcōŋ, K dərə(d^oə)
42. * ñī, ñīr *to sit*, M -ñīm, K -dā-
43. * ŋə *louse*, M -kit
44. * ŋre *egg*, K θii
45. * ŋre, ŋrer *to sing, dance*, M -kitec, K -θe-
46. * ŋri-re *small*, M -ktōŋnāŋ, K -riore *child*
47. * pa *I*, K -wa- *me, mine*
48. * par *foot*, M -pata?, K waa
49. * pat *anteater*, K wari(ri)
50. * pī *tree, firewood*, M mihim, mīm
51. * pri *path*, M pitahat, pitat, K rii
52. * pi-ci, -cit *one*, M picet
53. * rā *flower*, M -ta? *fruit*, K ra *fruit*
54. * rəp *dog*, K (ijə)rə
55. * ri *long*, M -toc, K irehe
56. * ta, tam *third person*, K d^oabī *to him*
57. * tə, tər *to fly*, M -to(paha?), K -ə-
58. * tu, tum *belly*, M -tec, K -wo-
59. * tik *black*, M -ññ
60. * twəm *fat, grease*, M -top, K wee
61. * za-ra *wing, feather*, K θa *feather ornament*
62. * za-re *root*, M (-ñip)catit, K (iru-)d^oi
63. * zaz-kwa *mouth*, M -ññ-koc
64. * zi *bone*, K d^oi
65. * zici *name*, M -cicet(?ac)
66. * zo, zoc *leaf*, M cic
67. * zi *seed*, K id^oi

Although the data are insufficient for a thorough reconstruction of Proto Macro-Jê, it is nevertheless possible to trace some of the major developments in the historical phonology.

A basic feature of Jê structure which apparently dates back to the Proto Macro-Jê horizon is the existence of parallel series of voiceless stops /p t c k/ and of nasals /m n ñ ŋ/.⁵ These are quite regularly

⁵ Members of the 'nasal' series are realized variously as nasal continuants, pre-nasalized or post-nasalized voiced stops, or even as simple voiced stops, depending on the specific environment and on the language involved. These are sometimes symbolized, as in the present Karajá data, by the normal voiced stop symbols.

reflected, especially in nonfinal position, in the languages under consideration. The evidence suggests that they have descended generally unchanged to Proto Jê but have undergone various types of change in both Maxakalí and Karajá. In Maxakalí the two series have become realigned so that, with few exceptions, both nonfinal */p t c k/ and nonfinal */m n ñ ŋ/ appear as /m n ñ ŋ/ contiguous to a nasal vowel and as /p t c k/ in other environments. At the same time, velars have undergone a split which has resulted in palatal reflexes, /c/ and /ñ/, under conditions that can probably be traced to the presence of a following front or central vowel at some stage in the history of the language. Final consonants have undergone other kinds of change, so that it is impossible on the basis of the present data to trace their development with any degree of certainty. Some final consonants, however, exhibit the same reflexes as nonfinal consonants.

The voiceless stop series and the nasal series are partly retained in Karajá. In most cases final consonants have been lost, although a few Karajá forms show the retention in non-final position of consonants that are reconstructed in final position for Proto Jê. Apart from final zeroing, */p/ is reflected as /w/ in prevocalic position and as zero preceding /r/. The alveolar stop is reflected variously as /r/, /d^ç/, or zero under conditions that are not clear. A single example indicates Karajá /č/ as a reflex of */c/. The velar stop is retained in a number of items of women's speech but is reflected as zero in most items of men's speech. The Proto Macro-Jê series */m n ñ ŋ/ is reflected in Karajá nonfinal position as /b r d zero/. It should be noted, however, that examples are lacking for Karajá reflexes of prevocalic */ŋ/.

In addition to the voiceless stop and the nasal series, Proto Macro-Jê probably had an */r/ phoneme as well as a phoneme corresponding to Proto Jê */z/. Except for some final occurrences, */r/ is reflected in Max-

akalí as /n/ when it is contiguous to a nasal vowel and as /t/ in other environments. In Karajá the phoneme has undergone a split, in which it is retained as /r/ in some environments and becomes /θ/ in other environments. The change to /θ/ was probably conditioned by a following front vowel, although the present evidence is not entirely clear. The Proto Jê phoneme */z/ is reflected in the Jê languages by a wide variety of sounds and its phonetic characteristics are unknown. In non-final position it corresponds quite regularly to Maxakalí /c/, or to /ñ/ when it is contiguous to a nasal vowel. In Karajá it corresponds regularly to the implosive /d^ç/.

Consonantal sound correspondences, except for irregularities and loss in final position, are summarized below. Numbers refer to the cognate sets previously listed.

PROTO JÊ		MAXAKALÍ
* p	p	48, 51, 52
	m	50
* t	t	52, 57, 58, 60
	n	40, 59
* c	c	2, 9, 28, 52, 65, 66
	ñ	8
* k	k	5, 8, 12, 15, 16, 19, 20, 24, 25, 63
	c	7, 9-11, 13, 18, 22
	ŋ	23
	ñ	8, 59
* m	p	13, 27, 32, 60
	m	28, 30, 31
* n	t	34, 37
* ñ	ñ	30, 38, 40-42
* ŋ	k	43, 45, 46
	ñ	5
* r	t	16, 18, 19, 22, 25, 32, 45, 46, 48, 51,
		53, 55, 62
	n	20, 46
* z	c	6, 24, 62, 65, 66
	ñ	63
PROTO JÊ		KARAJÁ
* p	w	47-49
	zero	51
* t	r	41, 49
	d ^ç	33, 56
	zero	57, 58, 60
* c	č	2
* k	k	3, 7, 14, 21, 22
	zero	4, 13, 15, 17-19, 22, 25, 39
* m	b	26, 29, 32, 33, 56

* n	r	35, 36
* ñ	d	29, 38, 39, 41, 42
* ŋ	zero	44-46
* r	r	19, 22, 25, 32, 35, 46, 51, 53-55
	θ	4, 18, 39, 44, 45, 61
* z	d ^h	6, 24, 62, 64, 67

In addition to the consonants covered in the above tabulation, Proto Jê */w/ has been reconstructed with a limited distribution. One cognate set (60) shows a corresponding /w/ in Karajá, while the remaining examples of Proto Jê */wV/ correspond to Maxakalí /o/ and to Karajá /u/ (2, 8, 25, 60, 63).

Vowel correspondences are considerably less regular than are consonantal correspondences, and few conclusions can be reached regarding vowel developments. It can be noted, however, that Maxakalí sequences of the type V[?]V (13, 30) and VhV (25, 32, 34, 50, 51) and Karajá sequences of the type VhV (14, 15, 55) correspond regularly to Proto Jê single vowels.

2. It is not possible on the basis of presently available evidence to arrive at a definitive classification of the Macro-Jê languages. The following is offered as a tentative outline in which are indicated those languages that can be classified with some degree of confidence, together with various lesser-known languages designated as unclassified.

I. The Jê family has the following subfamilies: A. Kaingangan subfamily (Several Kaingang dialects and Xokleng); B. Central Jê subfamily (Several Xavante dialects, Xerente, and extinct Akroá and Xakriabá); C. Northwest Jê subfamily (1. East Timbira language: Canela, Krinkatí, Gavião, Krahó, nearly extinct Krenjê, and several extinct local dialects; 2. Apinajé-Kayapó language: Apinajé and several Northern Kayapó dialects; 3. Suyá language).

Southern Kayapó is an unclassified Jê language.

II. The Maxakalí family consists of several Maxakalí dialects, all extinct except one; and Pataxó and Malalí, both of which are extinct or nearly so.

III. The Karajá family contains Karajá proper, Javahé, and Xamboiá, probably all dialects of a single language. Unclassified Macro-Jê languages are: Jeikó, Ofayé, Kamakanian (including Kamakan, Kutaxó, Masakará, and related languages or dialects), Purían (including Purí, Coroado, and Koropó), Botocudo (including several dialects). All of these are either extinct or on the verge of extinction.

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.

Languages listed as unclassified Macro-Jê are included largely on an impressionistic basis. Jeikó has long been extinct and is known only through a fragmentary word list collected by Martius early in the last century.⁶ It is placed in the Jê family in most classifications on the basis of some obvious similarities to the Jê languages. The most likely cognates in the data are: Proto Jê *krā, Jeikó *grangblá* *head*; Proto Jê *-mu Jeikó u(1)epu *to see*; Proto Jê *nā, Jeikó *nā* *mother*; Proto Jê *ñōtō, Jeikó *aenettá* *tongue*; Proto Jê *pa, Jeikó *aepang* *arm*; Proto Jê *par, Jeikó *aepahno* *foot*; and Proto Jê *zazkwa, Jeikó *aingko* *mouth*. Most comparable Jeikó items, however, show no similarity to reconstructed Proto Jê forms, and it is probable that the language does not belong in the Jê family but rather elsewhere in the Macro-Jê stock.

Ofayé has generally been placed outside of the Macro-Jê stock. Until recently very little data were available for this language formerly spoken in southern Mato Grosso. A few years ago considerable data were ob-

⁶ Carl Friederich Phil. v. Martius, *Beiträge zur Ethnographie und Sprachenkunde Amerikas zumal Brasiliens*, Vol. II, Leipzig, 1867.

tained by Sarah Gudschinsky from one of the last survivors of the tribe, and evidence presented in an unpublished paper shows this language to be related to Jê.⁷

In colonial times a group of non-Tupi tribes occupied a large area of what is now the state of Bahia as well as northern Minas Gerais and Espírito Santo. Loukotka has examined the available linguistic data for a number of these tribes and has concluded that they are related to Jê.⁸ He groups the languages into four families (corresponding to Maxakalian, Kamakanian, Purian, and Botocudo in the present listing), and cites some lexical similarities in support of their relationship to Jê. The similarities involve a number of items that can be reconstructed for Proto Jê and are sufficient to establish beyond reasonable doubt the fact that Kamakanian, Purian, and Botocudo do indeed belong within Macro-Jê. The exact classification of these languages will require a more thorough comparative study. In some cases this, unfortunately, may not be possible at this late date. The Purian languages have been extinct for some time and only fragmentary vocabularies exist. Languages of the Kamakanian group may also be extinct, but as late as in the 1930's a vocabulary was collected by Guérios.⁹ For Botocudo there probably remain a few individuals who can recall some of the language. In each case the written data are meager and the possibilities of field work either non-existent or fast becoming so.

Languages within the Jê family show at

⁷ Sarah C. Gudschinsky, *Ofayé-Xavante*, *Uma Língua Jê*, unpublished manuscript.

⁸ Chestmír Loukotka, *La Familia Lingüística Mašakali*, *Revista del Instituto de Etnología (Tucumán)* 2.21-47 (1931); *La Familia Lingüística Kamakan del Brasil*, *Revista del Instituto de Etnología (Tucumán)* 2.493-524 (1932); *La Familia Lingüística Coroado*, *Journal de la Société des Américanistes (Paris)* 29.157-214 (1937); *Les Indiens Botocudo et leur Langue*, *Lingua Posnaniensis* 5.112-135 (1955).

⁹ R. F. Mansur Guérios, *Estudos sôbre a língua Camacã*, *Arquivos do Museu Paranaense IV*, Curitiba, 1940.

least 40% shared cognates when compared one to another in contrast to a maximum of about 30% when compared to other Macro-Jê languages. Each family within the Macro-Jê stock is furthermore characterized by certain unique phonological developments. Maxakalian, for example, is characterized by the splitting of velars to yield palatals in some environments and the retention of velars in other environments. Karajá, on the other hand, is characterized by the double reflection of */r/.

Maxakalian includes, in addition to modern Maxakali spoken by an indigenous community in northeastern Minas Gerais, a number of languages and dialects now extinct or practically so. The internal classification of the Maxakalian family is a task that involves the same difficulties as those faced in the comparative study of Purian, Kamakanian, and Botocudo.

The Karajá family is represented by Javahé and Xamboiá in addition to Karajá proper, although data are available only for the latter. It is reported that there is a high degree of mutual intelligibility among the three tribes, but the exact nature and extent of dialect differences is unknown.

One Jê language, Southern Kayapó, is listed as unclassified. Although the language is now extinct, there are several word lists available. A careful examination of the data may result in the assigning of Southern Kayapó to one of the known Jê subfamilies, or as itself a coordinate subfamily.

Languages within each of the Jê subfamilies yield a minimum of 70% shared cognates when compared one with another, in contrast to a maximum of about 60% when compared with languages of other subfamilies. In addition, each subfamily is characterized by unique phonological developments. Kaingang, for example, shows /f/ or /ð/ as a reflex of */z/ in some initial positions. Central Jê has /h/ as a reflex of */k/ before a mid central vowel, and Northwest Jê shows /s/, /h/, or /ʔ/ as reflexes of */z/ in certain initial environments.

The internal structures of the Kaingangan and Central Jê subfamilies are not entirely clear. Further comparative study may clarify the picture and reveal whether diversity within each subfamily is merely dialect difference or whether distinct languages need to be recognized.¹⁰ Taven and Dorin, often listed as Kaingangan languages, probably represent Kaingang subgroups that no longer exist as distinct entities.

The Northwest Jê subfamily includes three distinct languages, two of which have numerous dialects. Lexical comparisons between dialects of the same language show a minimum of 80% shared cognates in contrast to a maximum of 75% in comparisons involving dialects of different languages. Reflexes of Proto Jê */z/ also serve to distinguish the three languages. East Timbira shows /ʔ/, Apinajé-Kayapó shows /h/, and Suya shows /s/ as initial reflexes of */z/ in certain environments.

Finally, it remains to note certain facts relating to possible wider relationships of the Macro-Jê languages—relationships of phylum order. There are in Brazil at least three languages or language families not included in Macro-Jê but which show quite striking similarities to the Macro-Jê languages. These are Borôro, the Tupi languages, and Fulniô.

Guérios has published evidence showing similarities between Borôro and certain Jê languages.¹¹ Some of the similarities, particularly those involving pronominal prefixes, are impressive, but the evidence needs to be expanded and systemized before it can be taken as proof of genetic relationship.

The resemblances between Tupi and Jê languages involve similarities in general

structure as well as many lexical similarities. Even a cursory examination of the available data reveals a good number of possible cognates. The following possibilities are drawn from my Proto Jê reconstructions and the Proto Tupi forms listed by Hanke, Swadesh, and Rodrigues:¹² Proto Jê *ma, Proto Tupi * pia *liver*; Proto Jê * mzen, Proto Tupi * men *husband*; Proto Jê * ηo, Proto Tupi * igi *water*; Proto Jê * ηo, Proto Tupi * ηkiv *louse*; Proto Jê * pa *arm*, Proto Tupi * po *hand*; Proto Jê * par, Proto Tupi * pi *foot*. Comparisons with individual Tupi languages reveal further possibilities: Proto Jê * krã, krãñ, Guajajara kaŋ- *head*; Proto Jê * ku, kur, Guajajara -ʔu *to eat*; Proto Jê * pri, Guajajara pe *path*; Proto Jê * pici, picit, Guajajara piteci *one*.

Fulniô is listed by Greenberg as an unclassified Macro-Jê language, but to my knowledge no evidence for this supposed relationship has been published. With a good amount of Fulniô data now becoming available it may be possible to demonstrate such a relationship through careful comparative techniques.¹³ The similarities in pronominal prefixes are quite striking. Both Jê and Fulniô have first person *i-* and second person *a-* prefixes. In addition, there are numerous other possible cognates, including some that suggest a correspondence between Proto Jê */p/ and Fulniô /f/: Proto Jê * prə, Fulniô fêlôwā *ashes*; Proto Jê * par, Fulniô fehe *foot*; Proto Jê * pika, Fulniô f'è'ā *earth*.

If these suggested relationships can be demonstrated, existing classifications of Brazilian languages will have to be radically revised. This possibility, however, must await more extensive reconstructions within the Macro-Jê and Tupi groups, as well as a closer examination of the Borôro and Fulniô data.

¹² W. Hanke, M. Swadesh, and A. Rodrigues, *Notas de Fonologia Mekens, Miscellanea Paul Rivet*, pp 187-217, Mexico, 1958.

¹³ Fulniô data are from Douglas Meland of the Summer Institute of Linguistics.

¹⁰ A preliminary study of Kaingangan dialects is found in Ursula Wiesemann, *Notas sobre Proto-Kaingang: Um Estudo de Quatro Dialetos*, to be published.

¹¹ R. F. Mansur Guérios, *O Nexo Lingüístico Bororo-Merrime-Caiapó*, *Revista do Circulo de Estudos 'Bandeirantes'*, Tomo 2º, No. 1, Curitiba, 1939.