

Lexical similarity between 2 Northern Nambiquara languages: a Mamaindé / Latundé comparative word list (including some Negaroté data as well)

*Dave Eberhard
Feb, 2005*

The following data is based on a dialect survey the author and a colleague were able to conduct during a field trip to both the Mamaindé and Latundé areas in December, 2004. The revised Swadesh word list used here was borrowed from Price (1978. The Nambiquara Linguistic Family. *Anthropological Linguistics*. 20: 23). This list adds some lexemes and deletes others in a manner that reflects Nambiquara culture and language considerations. In the list below, MD = Mamaindé, LT = Latundé, and NG = Negaroté. Latundé data are from the CD accompanying Telles (2002. *Fonologia e gramática Latundé/Lakondé*. Amsterdam: Vrije Universiteit). Mamaindé and Negaroté morphemes are from my own data. I omitted all items which were missing from one or both of the MD or LT lexicons, or which were impossible to find in the LT lexicon due to limitations on searching a PDF document for strings with special characters in them. The Negaroté column is mostly blank since my source of data on Negaroté comes from sparse entries in my database resulting from limited contact with Negaroté speakers. I would usually only record a Negaroté word in my Mamaindé lexicon if I noticed a word that seemed to diverge from the Mamaindé form. (My experience has been that the vast majority of the words are identical in Mamaindé and Negaroté and they usually have no problem understanding each other. There continues to be a lot of intermarriage between the two dialects.) Those spaces that are blank in this word list could then either be identical to the Mamaindé form, or they could be different but simply missing from my lexicon. Therefore, because the database used as a source for this wordlist was not meant to be a database of Negaroté, the wordlist should not be used as a source for comparing cognates between Negaroté and the other languages. This word list is only meant to compare Mamaindé and Latundé.

The presence of a term in a given row does not imply that it is the only possible term in that language with that same meaning. If there was more than one lexeme which presented the same meaning (synonyms), I chose the forms which were most clearly cognate in the three languages.

Note that almost all of the forms are cited as root morphemes only, regardless of whether or not the root is a bound morpheme. Typically, in the Nambiquara languages non-verbal roots can usually be uttered with minimal morphology, but verbal roots will always require a full suffix string. Leaving off most affixes was done to highlight the similarities and differences between single morphemes. The few forms which include affixes here are those which require the extra morphology to arrive at the indicated meaning. Such affix boundaries are indicated by hyphens in the Mamaindé data (the data from the other

languages did not include indications of morpheme breaks). Transcription is phonemic, and thus the reader must keep in mind that a number of phonological alternations occur in the surface forms of the data below, particularly in the coda position and in cases which include morpheme boundaries. Although the original Mamaindé data employs diphthongs, I'm utilizing the glides here (used by Telles in Latundé) in the place of diphthongs so as to maintain a uniform transcription among all the languages. Tone and stress are not marked. Shaded pairs are those which are *not* obvious cognates, or whose relationship is not yet clear.

<i>Portuguese</i>	<i>English</i>	<i>MD</i>	<i>LT</i>	<i>NG</i>
1. agora	now	wejn	wen	
2. água	water	nahon	nahoh	
3. amarrar	tie	tajs	koton	
4. andar	walk	ʔaj	ʔajh	
5. anta	tapir	nũnʔ-ta	nũntah	leun
6. aquele	that	hāj	hāj	
7. aqui	here	sejn	sen	
8. arco	bow	hukʔ	huʔ	
9. bater	hit	sun	sũn	
10. beber	drink	nã	nah	
11. bicho	animal	nũnʔ	nũn	
12. boca	mouth	ju	ju	
13. branco	white	hãn	hãn	
14. cabeça	head	na-kanĩn	najnkinĩn	
15. cabelo	hair	na-k ^h is	najnkih	
16. caçar	hunt/kill	ãn	ãn	
17. cair	fall	ta	tə	
18. caminho	road	tejh	teh	
19. campo	savannah	halo	haloh	
20. cantar	sing	hajn	haj	
21. capim	grass	sis	si	
22. carne	meat	sɨn	sɨn	
23. casa	house	sih	sih	
24. casca	bark	kalo	koloh	
25. cavar	dig	ʔawh	wen	
26. cheirar	smell	un	nĩn	
27. cobra	snake	tejh	təh	
28. coçar	scratch	k ^h ɨn	kɨn	
29. comer	eat	wɨ	wɨ	
30. comprido	long	lo	loh	
31. coração	heart	jalak ^h i	nuhoka	
32. criança	child	wejt	wet	
33. dar	give	tanũ	tanũh	

34. deitar	lie	t̄a	ta	
35. dente	tooth	w̄l	w̄l	
36. derrubar	chop	taw	tawn	
37. dois	two	pana	pan	
38. dormir	sleep	ĩn?	ũn	ĩwn?
39. ele	he	hāj	hāja	
40. seu	his	na	hāj	
41. escutar	listen	na-kas	nakat	
42. eu	I	t̄ai	t̄aja	
43. meu	my	t̄a	t̄a	
44. faca	knife	j̄un	uli?kaloh	walih
45. falar	speak	sejt	set	
46. feijão	beans	kamat	kamat	
47. flecha	arrow	?aw?-ti	aw	
48. flor	flower	hayāwn?gi	hūn	
49. folha	leaf	i-t ^h ān?	isān	
50. frio	cold	l̄	li	
51. fumaça	smoke	sinsi N	sin v	
52. homem	man	ejn?ni	en	
53. lenha	firewood	nī?	ni	nīw?
54. limpar	clean	nik?	ni	
55. liso	smooth	t ^h l	sih	
56. lombriga	worm	jujuk	kujuju	
57. longe	far	un	un	
58. lua	moon	ejn?	ejn?	
59. macaco prego	monkey 'prego'	hos	hoh	
60. mandioca brava	manioc - poisonous	l̄jn?	lin	
61. mão	hand	hik?	hi	
62. marido	husband	wetaj	witaj	
63. mata	forest	ta?wejn	tawen	
64. milho	corn	kajat	kejat	
65. morder	bite	i	i	iw
66. mulher	woman	tejhni	tēh	
67. nariz	nose	jamī	juwita	
68. noite	night	kanah	kanah	
69. olhar	see	ew	e	
70. olho	eye	i-kan̄jn	inkin̄n	
71. onde	where	nahajtā	āhāte	
72. orelha	ear	na-sa?lan?	nahalannaw	
73. osso	bone	n̄ah	n̄ah	
74. ovo	egg	nā	n̄a	
75. pai	father	m̄ini	mi	
76. pé	foot	j̄uk	j̄u	

77. pedra	stone	t̥ahna-kanĩn	t̥ajnkinĩn	
78. pegar	hold	tu	tu	
79. peixe	fish	ʔajn	ajn	
80. pena	feather	ʔwejs	weh	
81. pesado	heavy	saʔte	sateh	
82. piolho	louse	kanãjhni	kanajn	
83. preto	black	t ^h ot ^h on	toh	
84. puxar/arrastar	pull and drag	tasin	sĩntu	
85. quando	when	nanĩʔ-hejn	ãnihen	
86. que	what	natoʔ	ãtot	
87. queimar	burn	mãn	mãn	
88. reto	straight	ta-kakãn	mukakãn	
89. rir	laugh	kãwn	kown	
90. seco	dry	kamĩk	kamih	
91. sentar	sit	j̥au	tajau	
92. sujo	dirty	jejejs	jen	
93. ter medo	fear	suh	sun	
94. torto	crooked	takoʔ-takon	takown	
95. um	one	kanaka	kanah	
96. urucum	urucum	waʔtuh	hotu	
97. velho	old	ten	t̥en	
98. verde	green	saʔte-saʔten	tete	
99. vermelho	red	wawn	wawn	
100. vir	come	w̥a	wa	
101. voar	fly	i	i	
102. você	you	w̥aj	w̥aja	
103. seu	your	w̥a	w̥a	
104. voltar	return	ʔnĩ	ni	ʔnĩw
105. raiz	root	n̥ajk	n̥ajh	
106. ser pequeno	small	aaʔ	sipipi	
107. apertar	squeeze	tapas	t̥an	
108. braço	arm	nũkʔ	nuʔ	
109. cabaça	gourd	wal̥os	woloh	
110. homem adulto	adult man	jahon	ihon	
111. índio/gente	indian/person	nakaj̥an	kih̥nĩn	
112. sol	sun	sũni	sũn	

Additional data not part of the word list

113. esperar	wait	wãha	mah	mãha
114. tucano	toucan	jalãn	janãn	janãn
115. escorpião	scorpion	sasek	teteʔ	
116. roubar	steal	w̥anʔtu	walalah	
117. cupim	temite	walan	komũn	
118. trocar	trade	joha	wawn	

119. neto	grandchild	sawis	suníʔ
120. urinar	urinate	saʔkēwnʔ	tiniʔ
121. jabuti	tortoise	k ^h onʔ	kon
122. onça	jaguar	janān	loh
123. pequi	pequi	ʔan	lonkah
124. sumir	disappear	wate	wateh
125. estar doente	be sick	nakanis	najnkini

The first 112 of these items are from the revised Swadish list. These are the random sample. (The last 13 were items I came up with on my own, skimming the Latundé lexicon just to see if I could actually find more non-cognates between MD and LT. Thus they are not random and are included here mainly for reference. They should not be counted in the percentage of cognates.)

Results of the Comparison of Mamaindé and Latundé (out of 112 total words)

Definite cognates: 100

Not definite cognates: 12

(at least not obvious cognates where the relationships or sound rules are clear. Half of these 12 non-cognate forms were suspect and could possibly be considered as cognates, even though they are not viewed as such in this study are. These suspect forms are: #67, #96, #98, #99, #107, and #111.)

Cognate ratio between Mamaindé and Latundé: 89%

(this is a conservative estimate since the ratio would obviously be higher if the suspect forms were included)

The cognate ratio above shows that these two languages are quite close lexically. However, intelligibility tests that we conducted using the recordings of two Mamaindé texts indicated that intelligibility of Mamaindé speech among the Latundé is between 50% and 73% for those with no previous exposure to Mamaindé. This is much lower than the cognate ratio we have found here. This can be attributed to at least two factors:

1. **Different Phonology systems:** both the lexicons above are written in a phonemic transcription, without giving any indication as to the actual surface forms. I know from Stellas book that some of the complex phonology found in MD is absent in LT, thereby creating a considerable difference in pronunciation between the two – similar to the Spanish/Portuguese relationship.
2. **Roots vs. Affixes:** another factor is that there are few verbal affixes in the list above – it consists mostly of roots. But in these languages the verbal affixes can often constitute the majority of the sentence. So we aren't getting a clear picture of how cognate the affixes are between them, and in my short exposure to LT, it

seemed that it was in the verbal morphology that the languages differed the most, thus making comprehension more difficult. Of course, we would have to actually do a formal comparison of just the affixes across the two languages in order to get a clear indication of the distance between them at the affix level.