AMERINDIA n°35, 2011

Prosodic distinctions between the varieties of the Upper Xingu Carib language: results of an acoustic analysis

Glauber Romling da SILVA & Bruna FRANCHETTO

UFRJ, CNPq

1. Introduction: the Upper-Xingu Carib language and its varieties

The Carib subsystem of Upper Xingu consists of four local groups: the Kuikuro (four villages, with a fifth one being formed), the Matipu and Nahukwá (who live together in three villages), and the Kalapalo (two villages). All these groups speak a language which belongs to one of the two meridional branches of the Carib family (Meira & Franchetto 2005), and which nowadays presents two main varieties: one spoken by the Kuikuro and the younger Matipu generations, and the other spoken by the Kalapalo and the Nahukwá. Franchetto (2001) states that "we could establish a common origin of the Upper-Xingu Carib from which the first big division would have unfolded (Kalapalo/Nahukwá *vs*. Kuikuro/Matipu)."

These two varieties distinguish themselves by lexical as well as rhythmic differences. According to Franchetto (2001: 133), "in the Carib subsystem of the Culuene river, the interplay between socio-political identities of the local groups (*ótomo*) is based on distinct rhythmic and prosodic structures". Speakers express themselves metaphorically when talking about their linguistic identities. From a Kuikuro point of view (or

from whom is judging the other) we get the assumption of speaking 'straight' (*titage*) as opposed to speaking as the Kalapalo/Nahukwá do, which is 'in curves, bouncy, wavy' (*tühenkgegiko*) or 'backwards' (*inhukilü*) (Franchetto 1986; Fausto, Franchetto & Heckenberger 2008). In any case, the idea of 'straightness' as a way of speaking reveals a value judgment with regard to what it is not.

This paper is based on the proposal Franchetto (1997: 1) left for subsequent studies: "we seek, thus, a translation of this metalanguage in an analysis of the rhythmic structures of the two varieties." Our work now allows us to move forward in this investigation as corroborative acoustic analyses and new data have been included. We will try, somehow, to capture the reason for the native definition: I/we, 'straight' speech; the other, 'curved' speech.

The aims of this article are: (i) to establish the acoustic correlates that determine the position of stress in the varieties Kuikuro (KK) and Kalapalo (KP); (ii) to describe the stress patterns of both varieties and observe in which way they contrast; (iii) to offer an explanation to the metaphors put forward by the native speakers in light of the patterns observed.

2. Theoretical-methodological framework

Kager states the following regarding acoustic correlates for the perception of stress:

> Although the mental reality of prominence is undisputed, unambiguous phonetic correlate has not yet been discovered. Prominent syllables are potentially capable of bearing pitch movements with a strong perceptual load. They also tend to be of longer duration, as well as of a higher intensity, but both of the latter factors are usually subordinated to pitch (Kager 1995: 367).

Hayes defines 'culminativity' as that which makes us perceive just one syllable as the strongest in a domain. The concept of culminativity is thus connected to the definitions of acoustic correlates of pitch, duration and intensity: "*The overall picture, then, is that culminativity may be a universal of stress systems, which is subject to parametric variation for the level at which it holds*" (Hayes 1995: 25).

In short, culminativity is the principle or the universal, while the correlates that render its perception are the parameters which vary cross-linguistically.

The three acoustic parameters for the perception of culminative stress pointed out by Kager and Hayes are defined as follows:

– Pitch or fundamental frequency (F0): corresponds to the number of vibrations in/of the vocal cords in a given time frame: it is measured in hertz (Hz) and its perception is described in terms of pitch melody (low or high);

– Intensity: concerns the amplitude of the sound wave; it is measured in decibels (dB) and its perception is described in terms of volume (high or low, loud or soft);

– Duration: articulation time of a given domain (phone, syllable, word, phrase, sentence); it can be measured in seconds or milliseconds (ms or sec) and its perception is related to length (long or short).

According to Fry (1955, 1958 *apud* Hayes 1995: 6), out of the three parameters described above, the intensity is the parameter that has the smallest effect on stress perception. Its status is intuitive, as the most natural correlate for stress, revealed even in the way we locate culminative stress, indicating it as the 'strongest' or 'most intense' in a given domain. The duration has an intermediary effect, while pitch has a more relevant effect.

In some natural languages, only one parameter may determine stress perception, whereas in others more parameters interact. In Portuguese, we may say that duration determines perception; in Finnish, as there is a contrast between long and short vowels, the duration only affects non-stressed syllables in case of emphasis (Carlson 1978 *apud* Hayes 1995: 7); in Shilluk, an African language with contrasting lexical stress and tone, stress results from independent features in each type of contrast, such as is the case for Ma'ya (Libman 2005: 47). Therefore, our aim is to answer the following question: which acoustic correlates determine stress perception in Kuikuro and in Kalapalo?

In Figure 1, there is a template of the program PRAAT, that serves as an example to measure the three aforementioned acoustic parameters[1]. According to the illustrative arrows, we can see in the upper part the arrow that indicates syllable duration; in yellow, we have intensity, measured in decibels; in blue, pitch, measured in hertz.

[1] PRAAT is a free software program developed by Boersma & Weemink from the University of Amsterdam which can run on a wide range of operating systems.

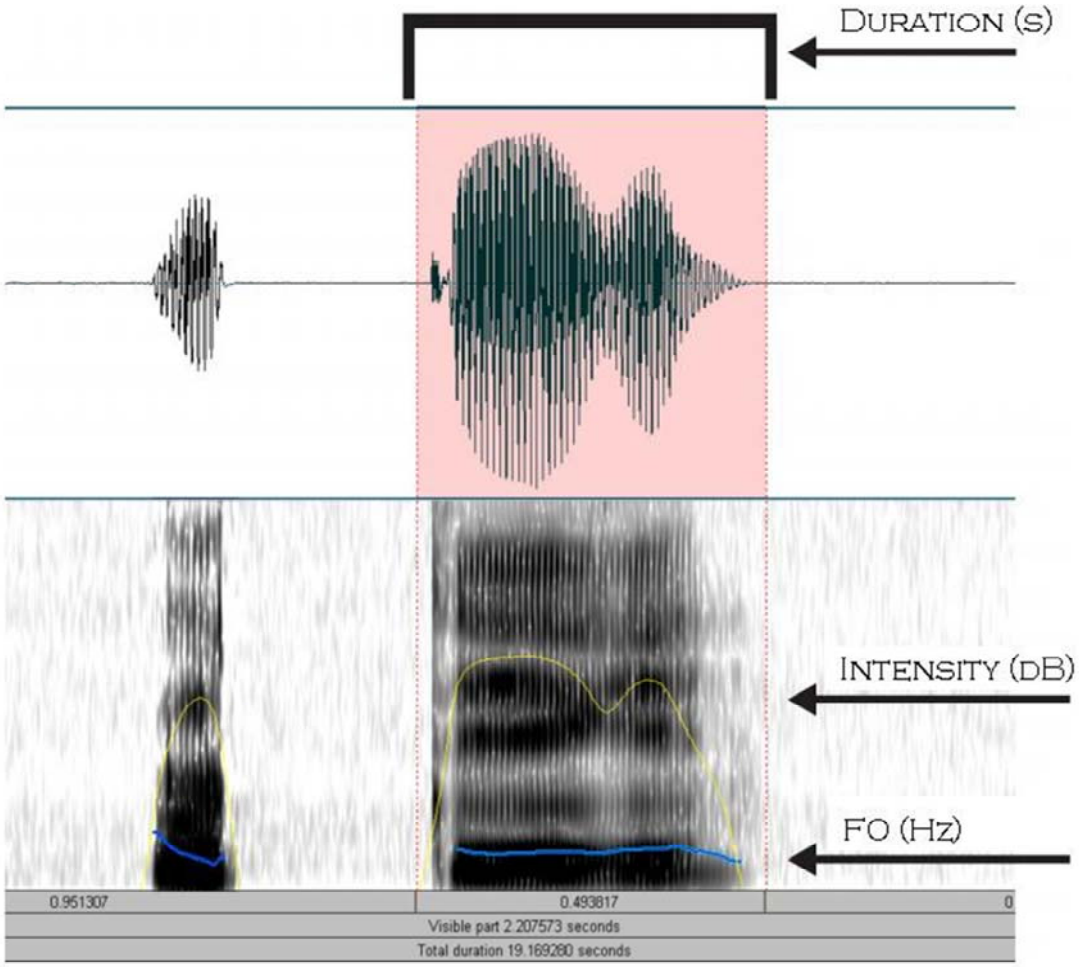

Figure 1: Template of a spectrogram in PRAAT

2.1. Corpus

The elicited sentences were constructed based on a target word. This word should vary in syntactic position and in number of syllables. The perception of culminative stress was considered in the domain of argument and head merging.

For each variety, we used a monosyllabic, a disyllabic, a trisyllabic, and a word containing more than three syllables.[2] All of them occurred in the following positions: (a) object of a transitive verb with a pronominal subject, (b) object of a transitive verb with a full nominal subject, (c) subject of an intransitive verb, (d) subject of a transitive verb with a full

[2] We are grateful to doctor Mara Santos, who kindly made available her Toolbox database. Without it, we would not be able to construct the sentences applied on this experiment.

nominal object. Here are the elicited structures followed by an example. The target word here is *ü* 'ax'.

a. Object [full noun] Transitive Verb Subject [pronoun]:

(1) *ü* *ingi-lü* *i-heke*[3]
ax see-PNCT 3-ERG
he saw the ax

b. Subject [full noun] Object [full noun] Transitive Verb:

(2) *kangamuke-heke* *ü* *ingi-lü*
child-ERG ax see-PNCT
(the) child saw the ax

c. Subject [full noun] Intransitive verb:

(3) *ü* *hugi-lü*
ax fall-PNCT
(the) ax fell

d. Object [full noun] Transitive Verb Subject [full noun]:

(4) *kangamuke* *ikenügü* *ü-heke*
child hurt ax-ERG
(the) ax hurt the child

In total, 16 sentences in Kuikuro and 16 sentences in Kalapalo were elicited. All of the sentences were elicited three times and only the second repetition was taken into consideration. For each variety we had a male and literate consultant in the age range between 25 and 30 years. The corpus was recorded using the Sound Forge program, at a 44KHz frequency with a 16 bits resolution. The microphone used was of the head set type, placed on the head of the consultant, 5 centimeters from the mouth.

2.2. Corpus analysis

For the analysis of the corpus we segmented the sentences into syllables. For each syllable we identified and considered only the stretch related to the vowel; we measured its duration, indicated the intensity peak in which we measured F0. Observe the following sentence:

---

[3] The examples in indigenous language are spelled according to the following conventions: *ü* (high central nonround vowel), *j* (voiced palatal consonant), *nh* (palatal nasal consonant), *ng* (velar nasal consonant), *g* (uvular flap). The abbreviations that are used for the interlinear gloss are the following: ERG= ergative case; PERF=perfective aspect; PNCT= punctual aspect; 3= third person.

(5) *kangamuke heke tahinga ingi-lü*
child ERG alligator see-PNCT
(the) child saw (the) alligator

The sentence is presented below in tables that were constructed with the values found in the Kuikuro and Kalapalo renditions, respectively; the syllables with stretches of white represent irrelevant values.

The values for F0, duration, and intensity will never be the same, even between speakers of the same variety, but we may notice, in an as of yet impressionistic manner, that the values for the acoustic correlates in Kuikuro and Kalapalo are quite different. This is the issue we will analyze in the following sections.

KUIKURO

	ka	ŋa	mu	kɛ	hɛ	ke	ˈta	hi	ŋa	ˈɲi	lɨ
F0 at I.P (hz)	–	105.66	103.06	118.72	–	102.64	108.1	96.24	113.84	96.15	97.06
Duration (sec.)	–	0.097	0.115	0.113	–	0.050	0.081	0.085	0.112	0.081	0.065
Intensity	–	71.62	70.16	72.83	–	72.62	75	66.15	73.84	69.65	64.94

KALAPALO

	ka	ŋa	mu	kɛ	hɛ	ke	ˈta	hi	ŋa	ˈɲi	lɨ
F0 at I.P (hz)		120.08	116.83	104.05		97.103	114.1	105.14	101.03	94.83	0
Duration (sec.)		0.110	0.038	0.129		0.050	0.099	0.060	0.106	0.069	0.055
Intensity		75.53	74.08	72.88		69.54	75.38	70.75	72.02	71.45	67.8

3. Analysis and results

3.1. Kuikuro

In Kuikuro, the acoustic correlate that is directly associated to stress is F0. The other parameters, duration (in seconds) and volume (in decibels), do not present any relevant regularities in their distributions. The syllable that is perceived as prominent is always the last of the internal argument which immediately precedes the verb. The association between the syllable perceived as prominent and the F0 peak may occur in two distinct manners. In the following section, the two perceived patterns and their respective generalizations are presented.

3.1.1 Pattern 1

Observe the behavior of two factors in the sentences (6), (7) and (8): stress position (marked by ') and F0 prominence.

(6) *tahinga* *ingi-lü* *i-heke*
alligator see-PNCT 3-ERG
he saw the alligator

	merge 1					*merge 2*		
	[argument			head	]	[argument	head	]
	[alligator		]	[see	]	[3]	[ERG	]
	ta	hi	'ŋa	ŋi	lɨ	i	hɛ	kɛ
F0 at I.P[4]. (hz)	104.33	110.95	117.43	***124.23***	106.49	–	113.4	
Duration (sec.)	0.077	0.073	0.119	0.117	0.057		0.129	0.021
Intensity at I.P(db)	74.19	68.02	73.35	69.43	71.49		74.34	54.19

(7) *kanga* *alamaki-lü*
fish fall-PNCT
(the) fish fell

	merge						
	[argument		head				]
	[fish	]	[fall				]
	ka	'ŋa	a	la	ma	ki	lɨ
F0 at I.P (hz)	117.1	125.3	-	***134.7***	102.8	whisper	
Duration (sec.)	0.069	0.133	-	0.052	0.090		
Intensity at I.P(db)	73.25	74.95	-	75.52	75.52		

(8) *tahinga* *alamaki-lü*
alligator fall-PNCT
(the) alligator fell

	merge							
	[argument			head				]
	[alligator		]	[fall				]
	ta	hi	'ŋa:	a	la	ma	ki	lɨ
F0 at P.I. (hz)	125.14	112.31	122.21		***125.96***	107.67	101.77	103.68
Duration (sec.)	0.085	0.101	0.080		0.038	0.092	0.043	0.074
Intensity at I.P(db)	77.34	71.96	74.67		73.36	75.46	62.3	59.12

[4] Intensity peak

In all of the examples, the syllable perceived as carrying the main stress is the last one of the argument, followed by the prominent F0 peak. The syllables which were unmarked for any of the parameters had an unperceivable acoustic signal. This pattern leads us to propose the following rule for Kuikuro:

(I) Rule 1 in Kuikuro: ['σ # **<u>σ</u>**], where 'σ stands for syllable perceived as prominent; **<u>σ</u>**, syllable with F0 peak.

3.1.2. Pattern 2

Observe below another pattern found in Kuikuro:

(9) *kangamuke heke ü ingi-lü*
child ERG ax see-PNCT
the child saw the ax

	merge 1					*merge 2*		
	[argument			head	]	[argument	head	]
	[child		]	[ERG	]	[ax]	[see	]
	ŋa	mu	'kɛ	hɛ	kɛ	'ɨj	ɲi	lɨ
F0 at I.P. (hz)	108.9	106.99	***<u>130.71</u>***		110.68	131.76	103.1	96.20
Duration (sec.)	0.115	0.187	0.168		0.158	0.142	0.109	0.081
Intensity at I.P(db)	71.95	69.25	74.18		73.11	77.16	71.11	66.91

In example (9), the syllable perceived as prominent is the one in which there is the prominent F0 peak. Therefore, in some constructions in Kuikuro, the two parameters can be combined. Thus, we have:

(II) Rule 2 in Kuikuro: [**<u>'σ</u>** # σ].

3.2. Kalapalo

For Kalapalo, in a relation of argument and head merge, we have an F0 curve which starts out high and decreases in a constant manner until the end of the merge, regardless of the number of syllables of the resulting phonological word. The association between the syllable that is perceived as prominent and the acoustic parameters which determine this prominence may occur in two ways. In the first case, only F0 is relevant; in the second, the other acoustic parameter, duration (in seconds) affects the perception of the stress. The two patterns are presented in the following sections.

3.2.1. Pattern 1

(10) *kangamuke heke tahinga ingi-lü*
child ERG alligator see-PNCT
the child saw the alligator

	merge 1					*merge 2*				
	[argument			head]		[argument			head]	
	[child		]	[ERG]		[alligator		]	[see]	
	ka	ŋa	mu	kɛ	hɛ ke	'ta	hi	ŋa	'ɲi	lɨ
F0 at P.I. (hz)		120.08	116.83	104.05	97.103	***114.1***	105.14	101.03	94.83	0
Duration (sec.)		0.110	0.038	0.129	0.050	0.099	0.060	0.106	0.069	0.055
Intensity at I.P(db)		75.53	74.08	72.88	69.54	75.38	70.75	72.02	71.45	67.8

In the example above, the syllable perceived as prominent is the one in which the F0 peak can be found. In argument and head merge (alligator#see), when the syllable is perceived as prominent, it is the initial syllable of this *merge*; F0 suffices to determine its perception.

3.2.2. Pattern 2

(11) *kanga alamaki-lü*
fish fall-PNCT
the fish fell

	merge							
	[argument		]	head				]
	[fish		]	[fall				]
	ka	'ŋa:		a	la	ma	ki	lɨ
F0 at I.P (hz)	119.7	***112.6***			101	82.27	0	sussuro
Duration (sec.)	0.076	***0.110***			0.056	0.093	0.039	
Intensity at I.P(db)	72.4	72.96			71.46	68.79	65.5	

In example (11) we see a different pattern than the one found before for Kalapalo. In this example, the syllable that is perceived as prominent is not the first of the merge. However, observing the parameter of duration, we may notice that the duration of the syllable perceived as prominent (0,110) is much longer than the duration of neighboring syllables (respectively, 0,076 and 0,056). In fact, the longer duration of the vowel may be attributed to a gemination resulting from the combination of the initial vowel of the verb (*alamakilü* 'fall') with the same final vowel of the preceding word (*kanga* 'fish'). However, below we have an example that may resolve this doubt:

(12) *kanga heke ihisü enge-pügü*
fish ERG fruit eat-PERF
(the) fish ate (the) fruit

	merge 1				*merge 2*					
	[argument		head	]	[argument			head		]
	[fish	]	[ERG	]	[fruit		]	[eat		]
	ka	'ŋa	he	kɛ	'i	hi	sɨ	ɲɛ	pɨ	gɨ
F0 at P.I. (hz)	115.6	***119.5***	112.3	95.33	119.4	118.4	104.4	96.76	0	-
Duration (sec.)	0.083	***0.123***	0.094	0.084	0.075	0.119	0.121	0.092	0.097	-
Intensity at I.P(db)	74.28	73.95	72.11	70.92	71.86	73.45	73.79	68.68	65.6	-

In (12) the syllable perceived as prominent is not the first of merge 1. The most prominent syllable presents a higher F0, but this is not the only acoustic parameter at stake: the duration of this syllable (0,123) also presents itself greatly superior to the value of the respective durations of its neighbors (0,083 e 0,094). This tells us that in the second pattern, when the initial syllable is not perceived as the most prominent, the non-initial syllable in which is perceived the prominence will present the duration parameter as its distinctive acoustic correlate . As such, we may say that duration is in a kind of 'complementary distribution' with regards to F0 in a rule that goes: perceive F0 as prominent in the beginning of *merge*; perceive duration as prominent in all other positions.

4. Conclusion

Based on the presented results, we may say that there is a distinction on two levels operating between the varieties of Kuikuro and Kalapalo. One distinction is acoustic in nature: Kuikuro uses only F0 for perception, whereas Kalapalo presents a complementary distribution between F0 and duration. Another difference is of a perceptual nature: Kuikuro maps the syllable perceived as prominent **within** the [argument#head] relation, while Kuikuto does so in **between** [argument#head] relations.

Therefore, in Kuikuro, perception-wise, the point of interest is the boundary between argument and head; in Kalapalo, the point of interest lies in the boundary between [argument#head] merges. Below there is a summary of the patterns that were found:

	Kuikuro	Kalapalo
Pattern 1	...[...'σ # **σ**...][...'σ # **σ**...]...	...[**'σ** ...# σ...][**'σ** ...# σ...]...
Pattern 2	...[...**'σ** # σ...][...**'σ** # σ...]...	...[...**'σ** ...# σ...][...**'σ** ...# σ...]...

As mentioned before, these differences are considered by speakers (at least by the Kuikuro) in terms of: 'we' with the 'straight' speech, versus 'the others' with the 'non-straight' speech. In the facts examined here, such metalinguistically highlighted differences were shown to correlate with empirical facts; it is interesting to observe that it is no coincidence that the perception of prosodic profiles is translated into metaphors that speak of movement, highlighting temporality and trajectories (be they straight or curved). Perhaps there is, in the domain of Upper-Xingu variants of Karib (and possibly in Upper-Xingu society as a whole), a metalanguage applying equally to speech, song, dance, drawing or graphic design. Nonetheless, by saying that speaking in a 'straight line' is good, while speaking 'in curves' is not, one adds a notion of 'straightness' which carries a positive value judgment against all that is 'deviant'. This is not surprising: it is a fact that native speakers of any language always observe the speech of another variety as 'singsongy' or 'with an accent' compared to the 'accentless' speech of whoever makes the judgment.

References

CARLSON, Lauri

1978 Word stress in Finnish. Unpublished manuscript. Massachusetts Institute of Technology.

FRANCHETTO, Bruna

1986 *Falar Kuikúro. Estudo etnolinguístico de um grupo caríbe do Alto Xingu*. Doctorate Dissertation, National Museum, UFRJ.

1997 Prosody and Dialect Distinctions in the Upper Xingu Carib Language. *49° Congreso Internacional de Americanistas,* Symposium "Lenguas Indígenas de las Tierras Bajas de América del Sur", PUCE, Quito (Ecuador), July 7-11.

2001 Línguas e História no Alto Xingu. *Os Povos do Alto Xingu – História e Cultura*, Franchetto B. & Heckenberger M. J. (orgs), 111-156. Rio de Janeiro: Editora da UFRJ.

FRY, Dennis
1955 Duration and intensity as physical correlates of linguistic stress. *Journal of the Acoustical Society of America* 27: 765-768.
1958 Experiments in the perception of stress. *Language and Speech* 1: 126-152.

MEIRA, Sergio & FRANCHETTO, Bruna
2005 The Southern Cariban Languages and the Cariban Family. *International Journal of American Linguistics* 71.2: 127-192.

FAUSTO, Carlos, FRANCHETTO, Bruna & HECKENBERGER, Michael J.
2008 Language, ritual and historical reconstruction: towards a linguistic, ethnographical and archaeological account of Upper Xingu Society. *Lessons from Documented Endangered Languages,* Harrison D. K., Rood D.S. & Dwyer A. (eds), 129-158. Amsterdam: John Benjamins.

HAYES, Bruce
1995 *The Metrical Stress Theory: Principles and Case Studies.* The University of Chicago Press.

KAGER, René
1995 The Metrical Theory of Word Stress. *The Handbook of Phonological Theory*, Goldsmith J. (org.), 367-402. Cambridge: Blackwell.

LIBMAN, Tatiana
2005 Acoustic Correlates of Stress in Shilluk (Nilo-Saharan). MA dissertation. The University of Edinburgh.