ORAL CONSONANT ACOUSTICS IN TIKÚNA (YURÍ-TIKÚNA)

Fernando O. de Carvalho

Graduate Program in Linguistics, Federal University of Rio de Janeiro (UFRJ), Brazil;
Laboratório de Línguas Indígenas (LaLI), University of Brasília (UnB), Brazil
fernaoorphao@ufrj.br

ABSTRACT

In this paper the results of an experimental investigation on the acoustic parameters employed in the implementation of a subset of the phonological consonants in the Tikúna language are presented. Special emphasis is placed on the discussion and presentation of variations that can be accounted for by reference to main stress placement in the language. It is shown how consonant duration, VOT and slope coefficients for Locus Equations vary as a function of the prosodic status of target syllables.

Keywords: Tikúna, stop consonants, prosody, locus Equations, VOT

1. INTRODUCTION

The Tikúna language (Yurí-Tikúna family; Carvalho [1]) is spoken by some 30.000 individuals in the Amazonian region encompassing the territories of Brazil, Peru and Colombia.

The part of the segmental inventory relevant to our study of the consonants of Tikúna includes the six plain stops /p t k b d ɡ/, the tap /ɾ/ and the two affricates /tʃ dʒ/. This is basically the inventory of oral consonants proposed by Soares [10, 11] and Rodríguez [9]. The language also has an approximant /w/ and a class of nasal consonants, though these are not included in the scope of the present paper.

One crucial aspect of this language's phonology concerns the placement of main word stress. Stress in Tikúna is morphologically determined, with a main stress mark being phonologically aligned with the leftmost syllable of a root. Furthermore, the main phonetic correlate of primary stress is an increased duration of the root's leftmost syllable (Soares [10]; Rodriguez [9]).

2. METHODS AND PROCEDURES

The data on which the present study are based consist of speech samples elicited from 4 adult speakers of Tikúna, 1 female (ME: 40 years old) and 3 male individuals (ET: 43 years old, JE: 38 years old and AT: 38 years old) all residents at the Ticuna de Santo Antônio reservation, Amazonas, Brazil. All four speakers are literate in Tikúna, having regular contact with written Tikúna material.

Nonsense words (phonologically possible in the language, but not attested) were employed and inserted into invariant carrier-sentences. Two sets of sentences were constructed: one in which the target syllables occur in stressed position (CV nonsense words; stressed condition) and the other in which they occur in an unstressed one (second syllable of a CVCV nonsense word; unstressed condition). The invariant carrier-sentence, meaning "This is a ____" was written in Tikúna orthography and presented in random order to the 4 subjects. Target syllables were CV combinations of the six phonological stops /p t k b d ɡ/ and the language's six oral vowels /a e i ɨ o u/. The nonsense words in the unstressed, CVCV condition, have [pa] as the invariable leftmost syllable, so that the condition is elsewhere in the paper referred to as paCV condition. The rhotic /ɾ/ was restricted to the unstressed condition, since it only occurs root-internally in Tikúna nouns.

The speakers' productions of the target sentences were recorded using a *Zoom H4* hand recorder with recording parameters set to: 48 kHz sampling frequency, 16 bits transmission rate and stereo mode.

The acoustic analyses were all carried using the *Praat* software for speech analysis. Stop consonant VOT was measured using the traditional procedures: the release burst is taken to be the reference point (zero) and its distance to the onset of periodic excitation of the following vowel is measured. This distance, computed over the time domain, gives the VOT value, in milliseconds. Positive VOT values are obtained when voicing follows the release burst (*Voice Lag* pattern) and negative values are obtained when voicing starts before consonantal release (*Voice Lead* pattern).

F2 values for the vowel's onset and steady portions were estimated using the LPC-based algorithm (*Burg*) for formant frequency estimation, set for 5 formants and with appropriate optimization of argument values as a function of the nature of the data. Visual inspection of spectrograms is as usual an additional aid to the estimation task. Optimization procedures included changing the *default* values for formant estimation as a function of the sex of the subject whose sample is under analysis.

Stop consonant duration was measured as the time domain distance between the beginning of the total signal attenuation created by the stop constriction (correlated with absence of formants in the spectrograms) and the release burst created by the constriction being undone. The same procedure was applied to measure the duration of the taps, with the proviso that here complete signal attenuation is seldom observed. Nevertheless, clear spectral 'gaps' were observed in spectrograms and were useful indexes of tap duration.

As for the affricates, the final transitional period characterized by fricative noise is included in the duration measurements as well. Data quantifying the percent presence of a clearly discernible release burst is also presented.

Descriptive statistical analyses and graph generation were done with *SPSS 14.0*.

3. RESULTS

3.1. Locus equations

For the description of particular CV sequences, Locus Equations (LE) were computed and these were taken as measures of coarticulation in these CV syllables (Tabain [13]; Krull [4, 5]). A LE is a linear equation (with its associated best-fit line) which describes the nearly linear relation obtained when two variables are plotted: the F2 value computed at the onset of the vowel, this marking the transition of the consonantal target gesture into the vowel target, and the F2 value obtained at the steady-state portion instantiating the vowel target.

Table 1: Slope values for each place of articulation as a function of the prosodic condition and pooled over both conditions.

Consoants	CV	paCV	Pooled
/p b/	0.873 R^2=0.963	0.860 R^2=0.950	0.862 R^2=0.955
/t d/	0.557 R^2=0.890	0.592 R^2=0.817	0.565 R^2=0.869
/k g/	1.057 R^2=0.987	1.113 R^2=0.979	0.901 R^2=0.982

For all three place of articulation categories there are higher slope values for the unstressed condition than for the stressed condition, except for the bilabials /p b/. This pattern reflects the prediction that coarticulation effects (that is, change in F2 onset locus as a function of F2 at the steady, vowel portion) are stronger in unstressed, hypoarticulated, than in prosodically stronger positions (Cole, Choi & Kim [2], Lindblom [3]). The slopes and the distribution of the data points are given in graphic form in figures 1-3, for each place of articulation category and independently of prosodic condition.

Both bilabial and velar stops show a clear segregation of the vowels in a 'front' group (/i e ɨ/) and a 'back' group (/a o u/) as indicated in the figures below:

Figure 1: Points given over F2 onset and F2 steady dimensions for CV syllables, with C=bilabial stop, independently of prosodic condition.

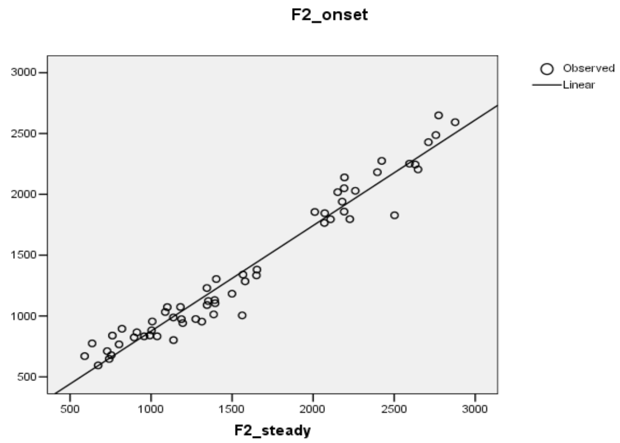


Figure 2: Points given over F2 onset and F2 steady dimensions for CV syllables, with C=velar stop, independently of prosodic condition.

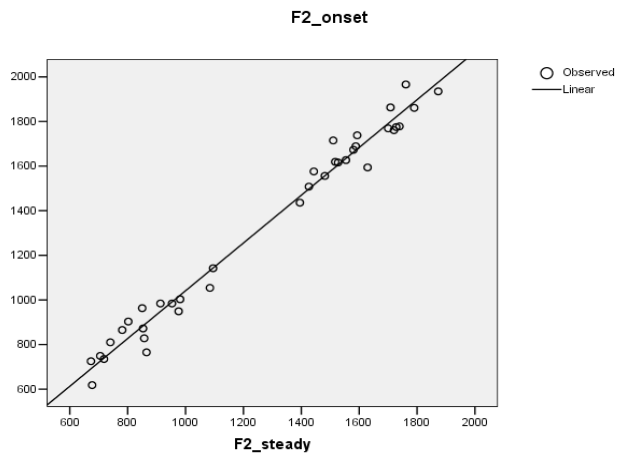
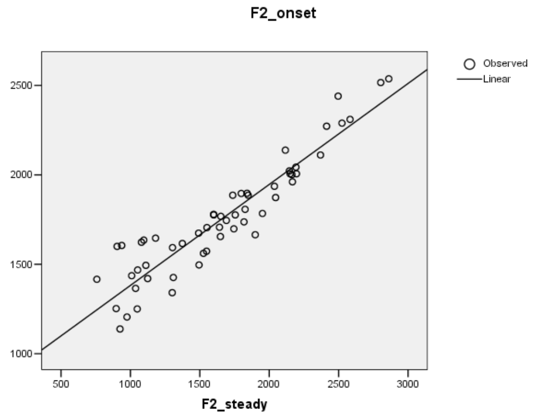


Figure 3: Points given over F2 onset and F2 steady dimensions for CV syllables, with C=coronal stop, independently of prosodic condition.



3.2. VOT and duration

Data on stop consonant VOT and duration are given below in tables 2-5:

Table 2: VOT and Duration for velar stops /g k/ as a function of prosodic condition.

Context		VOT (mean/S.D.)	Duration (mean/S.D.)
CV	/g/	-0.131 s / 0.0314	0.131 s / 0.0314
	/k/	0.034 s / 0.0096	0.158 s / 0.0241
paCV	/g/	-0.096 s / 0.0398	0.096 s / 0.0398
	/k/	0.041 s / 0.0177	0.144 s / 0.0195

Table 3: VOT and Duration for bilabial stops /b p/ as a function of prosodic condition.

Context		VOT (mean/S.D.)	Duration (mean/S.D.)
CV	/b/	-0.140 s / 0.0222	0.140 s / 0.0222
	/p/	0.006 s / 0.0016	0.191 s / 0.0444
paCV	/b/	-0.117 s / 0.0414	0.117 s / 0.0414
	/p/	0.008 s / 0.0025	0.158 s / 0.0236

Table 4: VOT and Duration for coronal stops /d t/ as a function of prosodic condition.

Context		VOT (mean/S.D.)	Duration (mean/S.D.)
CV	/d/	-0.140 s / 0.0211	0.140 s / 0.0211
	/t/	0.010 s / 0.0030	0.195 s / 0.0381
paCV	/d/	-0.101 s / 0.0231	0.101 s / 0.0231
	/t/	0.014 s / 0.0066	0.158 s / 0.0168

All voiced stops in our data are voiced throughout their duration (hence the same absolute values for VOT and duration in the tables above). Voiceless stops are consistently produced with positive VOT. The usual pattern in which velar stops have longer VOTs than coronal stops, with bilabials having the shorter VOT spans was observed in the Tikúna data.

Mean duration for all stops in larger in the stressed CV condition than in the unstressed paCV condition. The duration ratio between the stressed and unstressed conditions for each place of articulation and voicing specification are: /k g/: 1.09 and 1.36; /t d/: 1.23 and 1.38; /p b/: 1.20 and 1.19.

3.2.1. Affricates

For the affricates only VOT and duration data were taken. Information on place of articulation variation is also presented in the table below:

Table 5: Data for the affricates /dʒ tʃ/ as a function of prosodic condition.

Condition	Syllable	Mean VOT	Presence of Release Burst (% of tokens)	Duration of fricative segment
CV	/dʒ/	-0.126s	100 %	0.032 s
	/tʃ/	0.034s	94.4 %	0.060 s
paCV	/dʒ/	-0.035s	57.1 %	0.040 s
	/tʃ/	0.029s	100 %	0.063 s

As with the other voiced stops, the mean VOTs for the voiced affricates are all negative (i.e., the onset of voicing occurs before the release).

In the stressed, CV condition, a fricative /tʃ/ is sometimes implemented as a fricative [ʃ] *except* when followed by /ɨ/, in which case an affricate [tʃ] always is observed. The relatively low percentage of voiced affricates lacking a clear release burst in the unstressed (paCV) condition reflects the frequent implementation of these segments as an approximant [j], especially when /a/ follows.

3.2.2. The rhotic /ɾ/

We provide LE and duration data for the rhotic /ɾ/. The data presented in table 6 below show a comparison of the average duration of /ɾ/ in four other languages with the value here reported for Tikúna:

Table 6: Average duration of /ɾ/ in Tikúna compared to that of four other languages.

Language	Average constriction duration
Polish	20-25 msec.
Catalan	20-30 msec.
Spanish	20 msec.
Korean	20 msec.
Tikúna	21 msec.

The data above on the four other languages come from Jassem [3] (Polish), Recasens [8] (Catalan), Quillis [7] (Spanish) and Son [12] (Korean).

Another property of the Tikúna tap is that voicing was present in 100% of the tokens analyzed. This is not the case in languages such as Peruvian Spanish (Quillis [7]) or in Wari', a Chapakuran language of the Brazilian Amazon (MacEachern, Kern & Ladefoged [6]).

As a final observation, a LE slope value was also computed for the rhotic /ɾ/, yielding a value of b = 0.50 (R^2 = 0.823). The slope value is slightly lower than that found for the other coronal consonants, but is within the range of 0.40-0.60 usually found for taps (cf. Krull [5], Tabain [13]).

4. CONCLUDING REMARKS

The aim of this paper was to bring forth information on the phonetic structuring of a relatively little-studied Amazonian language, information that may be of typological and theoretical significance.

Data from Locus Equations as well as VOT and duration measurements were presented for the segments making the class of oral consonants in the Tikúna language (with the exception of /w/).

5. REFERENCES

[1] Carvalho, F.O. 2009. On the genetic kinship of the languages Yurí and Tikúna. *Revista Brasileira de Linguística Antropológica* 1(2), 247-268.

[2] Cole, J., Choi, H., Kim, H. 2003. Acoustic evidence for the effect of accent on CV coarticulation in radio news speech. *Proceedings of the 2003 Texas Linguistics Society Conference: Coarticulation in Speech Production and Perception.* Somerville, MA: Cascadilla Press, 62-72.

[3] Jassem, V. 1962. The acoustics of consonants. In Fry, D. (ed.), *Acoustic Phonetics: A Course of Basic Readings.* Cambridge University Press, 1976, 124-131.

[4] Krull, D. 1987. Second formant locus patterns as a measure of consonant-vowel coarticulation. *PERILUS* 5, 43-61.

[5] Krull, D. 1989. Second formant locus patterns and consonant-vowel coarticulation in spontaneous speech. *PERILUS* 10, 87-108.

[6] MacEachern, M., Kern, B., Ladefoged, P. 1996. Wari' phonetic structures. *Journal of Amazonian Languages* 1, 3-28.

[7] Quilis, A. 1981. *Fonética Acústica de la Lengua Española.* Biblioteca Románica Hispánica, Madrid.

[8] Recasens, D. 1991 On the production characteristics of apicoalveolar taps and trills. *Journal of Phonetics* 19, 267-280.

[9] Rodriguez, E.M. 1995. *Tonologia de la Lengua Ticuna.* CCELA: Universidad de los Andes, Colombia.

[10] Soares, M.F. 1992. *O Suprasegmental em Tikúna e a Teoria Fonológica*. Doctoral Dissertation, UNICAMP, Campinas, Brazil.

[11] Soares, M.F. 1995. Núcleo e coda: a sílaba em Tikúna. In Wetzels, L. (ed.), *Estudos Fonológicos de Línguas Indígenas Brasileiras*. Rio de Janeiro: UFRJ, 195-263.

[12] Son, M. 2008. Pitfalls of spectrogram readings of flaps. *Acoustics'08* Paris, 693-698.

[13] Tabain, M. 2000. Coarticulation in CV syllables: a comparison of locus equation and EPG data. *Journal of Phonetics* 28(2), 137-159.