

## A methodology for the study of rhythm in drummed forms of languages: application to Bora Manguaré of Amazon

Julien Meyer<sup>1,2</sup>, Laure Dentel<sup>1,2,3</sup>, Frank Seifart<sup>4</sup>

<sup>1</sup>Área de Linguística, Museu Goeldi, Ministério de Ciência, Tecnologia e Inovação, Belém, Pará, Brasil

<sup>2</sup>Sound Communication and environmental auditory Perception Research Group, Paris, France

<sup>3</sup>Engenharia Elétrica, Universidade Federal do Pará (UFPA), Belém, Pará, Brasil

<sup>4</sup>Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany

julmeyer.lab@gmail.com, lauredentel@ufpa.br, frank\_seifart@eva.mpg.de

### Abstract

This study presents a new methodology adapted to the analysis of word rhythmic cues in drummed forms of languages. The semi-automatic beat detection procedure applied to the Manguaré drummed form of the Bora language enabled us to measure inter-beat durations. These were found to correspond to Vowel-to-Vowel intervals (V-to-V) of the associated speech utterances and to differ as a function of the vowel duration and of the presence/absence of consonant(s) in the V-to-V cluster.

**Index Terms:** drummed language, rhythm, prosody, Bora language, semi-automatic detection, phoneme clustering.

### 1. Introduction

#### 1.1. General background

Drummed forms of spoken languages have been so far reported to exist in Africa, South America, Asia, and Oceania [1, 2, 3]. They consist in the emulation of phonological and prosodic features of spoken utterances by the means of drummed beats. They are devoted to long distance communication, exploiting the natural bio-acoustical properties of drummed signals for a good propagation of sound in natural environments. For example, drummed signals resist well to acoustic energy loss due to reverberation in forests because low pitch frequencies are not blocked by large vegetation obstacles. Another aspect is that the high amplitude of percussions overcome the ambient noise at greater distances than normal and shouted voice while their narrow frequency-band reduces the noise-masking effects.

These drummed imitations of speech must be distinguished from other categories of complex drummed signaling or musical systems sometimes also called 'drummed languages' but that are not based on spoken language. Indeed, some traditional signaling systems use repertoires of drummed codes with no iconic relationship to the sound structure of the locally spoken language. These constitute parallel communication systems confined only to drumming and are for example particularly often attested in Melanesian cultures [3, 4]. On the other hand, some musical drumming traditions use vocalized nonsense syllables (or vocables) in systematic ways to represent drummed sounds, mostly for mnemo-technical purposes. Some of these musical associations are based on an acoustic and perceptual mapping between vocable sounds and musical sounds, like in North Indian tabla [5], whereas others are purely symbolic, like

the familiar occidental solfège, in which the notes of the scale are represented by the syllables *do, re, mi, fa, sol, la, si*.

The present study deals only with the category of drummed beats related to speech for distance communication, i.e. in which the relationship between the signifier (the drummed signal) and what is signified (the speech utterance) is not purely symbolic like in codes, but based on a relation of physical similarity combining abridgment [1] and acoustic iconicity [2]. A large review on such drummed forms of languages was made by Sebeok and Umiker-Sebeok [2]. It provides information about 18 different languages including Ewe, Twi (Akan), Banen, Chin, and Bora, mostly by reprinting descriptions from the 19<sup>th</sup> and first half of 20<sup>th</sup> century. It showed that such abridgment systems are most frequently characterized by the reproduction of tonal contrasts and rhythmic elements of the base spoken utterance. **Tonal contrasts:** the majority of drum forms of languages are associated with tonal languages. In such cases phonological surface tone patterns are exactly rendered through different drum pitches. Register tones are represented by single beats, and contour tones are often rendered by a succession of two beats of different pitches. **Rhythm:** the parameters concerning rhythmic patterns have hardly been tackled in previous works, particularly inter-beat durations (IBDs) in words. So far, rhythm was mostly dealt with simple characterizations, asserting that the general rhythm of normal speech was mimicked by drummed sentences, with pauses between sentences and a beat corresponding to each syllable [6, 7]. However, few articles analyzed these aspects in detail. For example, Nketia [8] found an association between the relative lengths of inter-beat durations and syllable weight, in which closed syllables corresponded to long IBDs; open syllables followed by a voiced consonant or a vowel were associated to extra-short IBDs; and other syllables were drummed as simple short IBDs. On the other hand, Cloarec-Heiss [9] was the only one to link some aspects of drummed rhythm to classes of consonants as she found that liquid consonants [l, r, v] corresponded to shorter inter-beat durations than other consonants.

#### 1.2. The Manguaré drummed form and the Bora language

Bora Manguaré is one of the very few drummed signaling systems documented so far in South America, and it is the only one known to be directly linked to an Amerindian spoken language. Manguaré is composed of two hollow log drums of

different sizes. Each drum has two pitches (one on each side of the slit) and therefore the pair of drums has four pitches. Only two (one of each drum) are used for the speech mode of drumming. According to previous publications, Manguaré drums emulate formulaic sentences and each beat represents the tones of a syllable (Bora having two contrastive tones (H and L)) [10]. Relatively informal messages are produced with the Manguaré, for example to make public announcements, or for ‘calling messages’, which may be used on a daily basis, usually to ask someone to bring something or to come.

In the present study we explored how, besides encoding pitch, such drummed messages may also encode linguistic information through inter-beat durations. We developed an original method of semi automatic tracking to calculate such inter-beat durations. For such a study we computed each drum message as a function of its associated phonemic content. Bora language has six phonemic vowel positions in addition to a phonemic distinction of vowel length: [i, ī, u, ε, a, o]. Moreover, the Bora plain consonants [p, b, t, d, ts, tʃ, dz, dʒ, k, k<sup>h</sup>, k<sup>w</sup>, ʔ, β, r, m, n, h] almost all have palatal(ized) counterparts except [ts, tʃ, dz, dʒ]. The syllable structure of Bora is (C)V(C), with the restriction that only the glottal fricative (h) and stop (ʔ) may occur in coda position and only if the vowel is short. The four main syllable types are thus (C)V, (C)V̄, (C)Vh, and (C)Vʔ [11].

As far as we know, no other existing publication on drummed forms of languages provided such a statistical analysis of rhythm because none had sufficient data combined with an adapted methodology to look at this aspect.

## 2. Method

### 2.1. Materials

95 typical Manguaré drummed messages (mostly ‘calling messages’) were elicited from five different expert Bora Manguaré drummers. These messages were recorded and transcribed in ELAN [12] by Seifart together with a native speaker/drummer. This original data is archived and accessible online [13]. Each of these messages contained about 15 words and a total of about 60 drum beats. After a pre-analysis of this data, we selected the productions of the drummer which had the largest repertoire of messages, composed of 1806 word utterances and 4452 pairs of drum beats/intervals. However, due to the high rate of repetition in drummed messages, we identified only 197 different words in total, repeated non homogeneously (Figure 1)

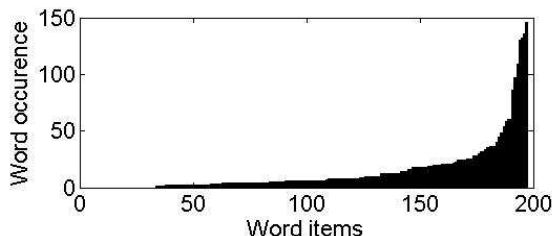


Figure 1 : Distribution of word occurrence. Each word was attributed here a number of word item.

The most frequent words corresponded to stereotyped calling formulas present in almost every messages. In reason of particular status and their very high weight in statistical results,

we decided to eliminate them. We therefore finally analyzed 2758 interbeat durations, with 1194 Low (L) tone beats and 1574 High (H) tone beats.

### 2.2. Design and procedure

After having exported from ELAN the initial and final time of each transcribed speech utterances, the inter-beat durations were automatically measured using software specially developed for the occasion.

#### 2.2.1. First step: phoneme clustering

We applied a method of phoneme clustering associated to each Manguaré sound file, after having coded consonants as C<sub>x</sub> (x=1 to 17, alphabetic order); and vowels as V<sub>y</sub> (y= 1 to 12, alphabetic order). Glottal stops (ʔ) and glottal fricatives (h) were for example coded as C6 and C7. Moreover, the palatal(ization) was identified and coded as P. Two alternative strategies were foreseen, that is, two competing hypotheses to test:

- Syllable clustering*: in this case the inter-beat duration would be associated to the syllable duration; phonemes would be grouped following the syllable segmentation of the word (Table 1, second row).
- V-to-V clustering*: in this case, the inter-beat duration would be associated to the phonemes present between two consecutive vowels. Each beat would be associated to the maximum amplitude of each vowel of the word (Table 1, third row).

Table 1: Illustration on a word of the two competing phoneme clustering hypotheses.

	k <sup>w</sup>	á	r	ʔ	ʔ	k <sup>h</sup>	o	ʔ	í	h	k <sup>h</sup>	j	a	k <sup>h</sup>	i
a)	CV	CV	CCV	CV	CCPV	CV									
b)		VC	VCC	VC	VCCP	VC									

Each type of clustering strategy of the inter-beat durations resulted in a different distribution of the involved phonemes of (see Table 2 and 3). As a result some types like VV didn't exist in the V-to-V distribution as there was no utterance with a long syllable followed by another syllable in the involved corpus.

Table 2: Number of clusters per syllable type (VV=V: and CC= Ch, hC, ʔC or Cʔ)

Syllable types	V	VV	CV	CVV	CCV	CCVV
Cluster Nb	542	37	1490	256	431	12

Table 3: Number of clusters when the phonemes are redistributed according to a V-to-V clustering strategy (VV=V: and CC= Ch, hC, ʔC or Cʔ)

V-to-V types	V	VV	VC	VVC	VCC	VVCC
Cluster Nb	232	0	1717	305	514	0

#### 2.2.2. Second step: automatic beat detection

We first applied a sound filtering using a beat synchronization algorithm that automatically detects beats in a recorded piece [14]. Next, we estimated the sound energy for each beat. Drum beat timings were detected as the maximum peak of amplitude of a beat. To ease the performance of automatic detection, various contextual thresholds were introduced such as: (i) a minimum

limit of energy peak which depended on the recording level and the drum power; (ii) a minimum limit of inter-beat duration, which depended on the drummer (to suppress sliding effects of the stick); (iii) and a maximum limit of the energy rate between two consecutive beats, which depended on all the preceding parameters, in order to suppress the strong energy peaks which were not drummed transients.

### 2.2.3. Third step: peak and cluster association

The detected beat timings were validated when the number of detected peaks of a Manguaré sound file matched with the number of phoneme clusters in the corresponding ELAN file (Figure 2, top and center). Each beat time was manually verified before performing the measure of inter-beat durations.

### 2.2.4. Fourth step: frequency estimation

Finally, the power spectral density was calculated around each peak time. The peak frequency was calculated as the frequency of the maximum power (Figure 2, center). We verified that the H/L tone value of the vowels matched respectively with 141 Hz (High pitched drum) or 94 Hz (Low pitched drum).

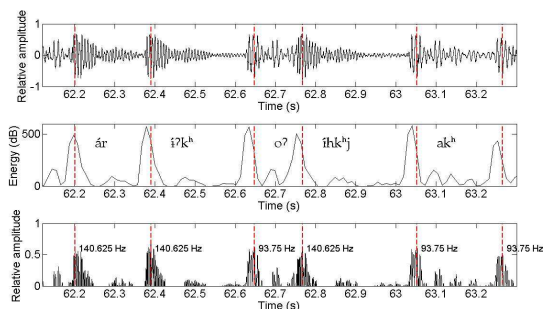


Figure 2: Drum sound file (top), associated phonetic transcription (center), energy estimation (center), and automatic beat filtering for beat detection (bottom)

## 3. Results

### 3.1. Inter-beat durations and V-to-V intervals

We have seen in 2.2.1 that *Syllable clustering* and *V-to-V clustering* strategies theoretically differ in terms of consonants associated to the beat-to-beat interval, but not in terms of vowels (cf. example Table 1). The results we found for inter-beat durations show that, effectively, the phoneme clusters associated to long vowels (VV) are drummed significantly longer than the phoneme clusters associated to short vowels, for both clustering strategies (see for example IBDs of V vs VV and CV vs. CVV in Figure 3, left; and IBDs of VC vs. VVC in Figure 3, right). However, they also show that, in the case of the V-to-V clustering, an influence of the presence of consonants was found. To illustrate that, the data presented on Figure 3 show that there exists a scale of IBDs that can be ordered as follows in the V-to-V clustering case:

$$IBD(V) < IBD(VC) < IBD(VVC) \sim IBD(VCC)$$

Indeed, the IBD difference between VVC and VCC is the only to be borderline significant in this case (one way anova:  $F(1,817)=4.59$ ;  $p=.0324$ ).

On the contrary, in the syllable-clustering case, the results presented on Figure 3 show that the onsets of syllables involved in an inter-beat interval either do not influence the IBD (we found a non significant difference between VV and CVV syllables ( $F(1, 291)=.3$ ; n.s.)), or influence it in a not logical way (we found that the IBD comparison between V and CV groups ( $F(1,2030)=74.85$ ;  $p<.0001$ ) and between V and CCV groups ( $F(1,971)=19.11$ ;  $p<.0001$ ) indicate that adding consonants to a vowel would significantly diminish the interval durations in syllable clustering condition).

As a consequence, these results converge to support the hypothesis that the drum peak corresponds to the vowel nucleus and does not encode the onset consonant of the syllable. Therefore, we have enough elements to argue that the inter-beat duration reflects an interval more closely related to the vowel-to-vowel interval rather than to the syllable. As a result, we chose the V-to-V phoneme clustering strategy for further analyzes of our data.

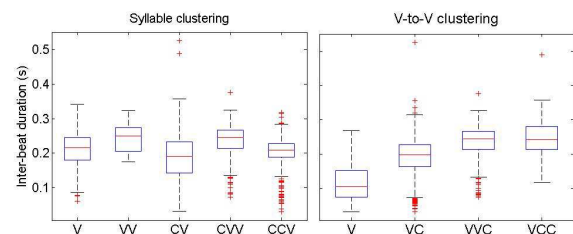


Figure 3: Distribution of inter-beat durations as a function of clustering strategy and cluster types. (VV=V: (long vowel) and CC= Ch, hC, ?C or C?)

On the basis of these results and to go deeper in the analysis, we looked for more factors influencing the inter-beat durations. For example palatalization, tone and vowel type of the first beat, tone and vowel type of the second beat, or consonant type.

### 3.2. Asymmetry in inter-beat durations depending of tonal configurations

The factor with the clearest impact on IBDs concerned the tone sequence of the two beats delimiting each inter-beat interval. Namely, the effect of the four possible tonal patterns LL, LH, HL, HH on IBDs was measured as statistically significant on V clusters ( $F(2,229)=77.16$ ;  $p<.0001$ ), on VC clusters ( $F(3,1716)=436.35$ ;  $p<.0001$ ), on VVC clusters ( $F(3,304)=38.97$ ;  $p<.0001$ ) and on VCC clusters ( $F(3,513)=58.98$ ;  $p<.0001$ ).

To take this effect into account we refined the presentation of the previous results (Figure 3) by classifying the IBDs as a function of tonal duplets (Figure 4). The LH pattern, with the greatest number of items (between  $n_{min}=68$  for VVC and  $n_{max}=418$  for VC) showed the shortest IBDs in any kind of V-to-V structure and it respected the duration patterns of Figure 3 (right). Next, HH (with  $n_{min}=32$  for VCC and  $n_{max}=381$  for CV) was found longer in duration but also with less compact IBDs. However, HH tonal configurations still respected the duration patterns of Figure 3 (right). Finally, LL and HL generally bore the greatest durations but HL was not enough present in V clusters and LL was not enough present in VVC clusters to compare their distribution across V-to-V types of Figure 3.

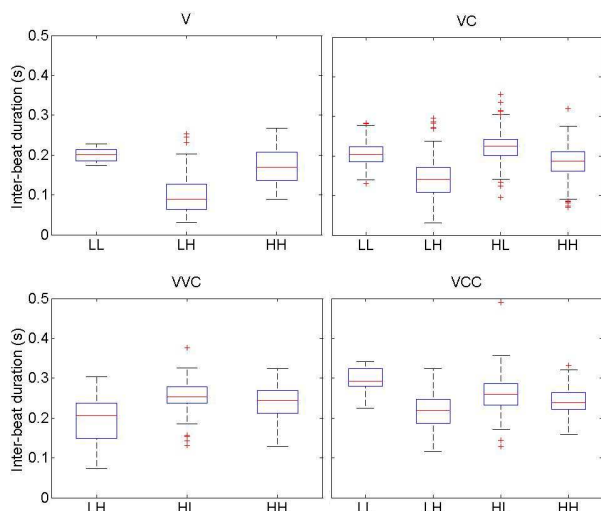


Figure 4: Distribution of inter-beat durations as a function of clustering strategy and cluster types. (VV=V: (long vowel) and CC= Ch, hC, ?C or C?).

#### 4. Discussion and conclusions

This study provided an original methodology to analyze the word rhythmic structure of drummed forms of languages. We applied this methodology to the drumming traditions of Bora language of Amazon and found various important results. First, we showed that the inter-beat durations played by the Bora drummers corresponded to V-to-V intervals rather than to syllables of the associated spoken utterances. Such a result is not totally surprising as some linguists have already underlined the important role of V-to-V clusters in the rhythmic structure of spoken languages [15, 16]. Moreover, we found that IBDs depended of the constituent phonemes of the V-to-V cluster, with IBDs of single vowels shorter than IBDs of VC clusters, themselves shorter than IBDs of long vowels and CCV clusters. Such rhythmical features at the word level must play an important role in the recognition of drummed messages as tone discriminates only relatively few lexical items in Bora language, but has many grammatical functions [11]. Interestingly, the asymmetry of IBDs that we found to depend on tonal configurations didn't break the IBD differences due to phoneme clusters. Some preliminary measures that we made on the spoken form of speech for V clusters seem to indicate that a dependency of V-to-V interval durations to tonal duplets also exists in spoken Bora. The systematic comparison of drummed IBDs to V-to-V spoken IBDs is therefore a promising further development of this research.

Another possible further development concerns individual consonants. We didn't develop here the distribution of IBDs as a function of consonant types because on the few cases with enough items in our data ([h, m, k<sup>w</sup>] for example) IBDs in LH configurations were not coherent with IBDs in HL configurations. This shows that the corpus of measured IBDs must be extended to analyze each Bora drummed consonant separately and eventually detect the influence of other factors on their realization.

#### 5. Acknowledgements

The authors would like to thank the Bora drummers and speakers, and particularly the curaca Manuel Ruiz Miveco. This research has been supported by the Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq (Pós Doutorado Junior Fellowship n°159747/2010-4; and Study Fellowship n°155079/2010-7). Data collection was supported by Volkswagen Foundation DoBeS grants II/ 80 110 and II/83 522.

#### 6. References

- [1] Stern, T., "Drum and whistle languages: an analysis of speech surrogates", *American Anthropologist*, 59: 487-506, 1957.
- [2] Sebeok, T.A. and Umiker-Sebeok, D.J. [Eds.] "Speech Surrogates: Drum and Whistle Systems. Approaches to semiotics" 23. The Hague: Mouton, 1976.
- [3] Niles, D., "Editor's Introduction to Zemp & Kaufmann, Kulele: Occasional Papers on Pacific Music and Dance", 4: 1-50, 2010.
- [4] Zemp, H. and Kaufmann, C., "Pour une transcription automatique des "langues Tambourinés" mélanésien", *L'Homme* 9: 38-88, 1969.
- [5] Patel, A.D., "Music, Language, and the Brain". NY: Oxford University Press, 2008.
- [6] Agawu, V. K., "African rhythm. a northern Ewe perspective". Cambridge: Cambridge University Press, 1995.
- [7] Carrington, J. F., "Talking Drums of Africa". London: Carey-Kingstate Press, 1949.
- [8] Nketia, J. H. K., "Drummed languages of Africa". in T. Sebeok and D.J. Umiker-Sebeok [Eds.]. *Speech surrogates: drum and whistle systems*. La Haye and Paris, 825-864, Mouton, 1976.
- [9] Cloarec-Heiss, F. "From natural language to drum language: An economical encoding procedure in Banda-Linda, Central African Republic". *Language diversity and cognitive representations*, in C. Fuchs and S. Robert [Eds.], 145-157, Amsterdam, Philadelphia: John Benjamins Publishing Company, 1999.
- [10] Thiesen, W., "The Bora signal drums". *Lore* 19: 101-103, 1969.
- [11] Seifart, F., "The structure and use of shape-based noun classes in Miraña (North West Amazon)", Nijmegen: Max Planck Institute for Psycholinguistics, 2005.
- [12] Brugman, H. and Russel, A., "Annotating Multimedia/ Multimodal resources with ELAN". *Proceedings of the 4th International Conference on Language Resources and Evaluation*, 2004. (<http://www.lat-mpi.eu/tools/elan/>).
- [13] Seifart, F., Doris F., Gasché J. and Echeverri J. A. [Eds.] "A multimedia documentation of the languages of the People of the Center", 2009. ([http://corpus1.mpi.nl/qfs1/mediaarchive/dobes\\_data/Center/Info/WelcomeToCenterPeople.html](http://corpus1.mpi.nl/qfs1/mediaarchive/dobes_data/Center/Info/WelcomeToCenterPeople.html)).
- [14] Cheng, K., Nazer, B. Uppuluri, J. and Verret., R., "Beat This: A Beat Synchronization Project", 2001 ([http://www.clear.rice.edu/elec301/Projects01/beat\\_sync/index.html](http://www.clear.rice.edu/elec301/Projects01/beat_sync/index.html)).
- [15] Marcus, S.M., "Acoustic determinants of perceptual-center (p-center) location", *Percept. Psychophys.* 30 (3): 247-256, 1981.
- [16] Pompino-Marschall, B., "The syllable as a prosodic unit and the so called P-centre effect". *Forschungsberichte der Institut für Phonetik und Sprachliche Kommunikation der Universität München* 29: 66-124, 1991