

Narrow focus in Brazilian Portuguese: spatial and temporal constraints

Luciana Lucente and Plínio A. Barbosa

Speech Prosody Studies Group / Department of Linguistics
Campinas State University, Brazil

luciana.unicamp@gmail.com; pabarbosa.unicampbr@gmail.com

Abstract

This paper presents the Brazilian Portuguese intonational annotation system, the ToBiPI system, for which the use of a set of labels integrates a dynamical annotation with some aspects of intonational phonology research paradigm, namely the intonational workshops. The dynamic annotation is based on a function-oriented analysis that takes into account some spatial and temporal constraints such as alignment and rate of falling and rising.

1. Introduction

Intonation is generally defined as the combination of tonal features and acoustic parameters, such as fundamental frequency (henceforth f_0). The term intonation is used to describe tonal features specifically, used by speakers to mark the pragmatic force of information in an utterance [6], which gave rise to the development of several intonational annotation systems based on f_0 descriptions. The term intonation is used interchangeably with the term prosody, which involves on their description temporal and dynamic features, as the alignment of intonation contour with syllables and vowels [6].

The definition of intonation adopted in this paper is that of an intonation event which can provide information to the listener about the prosodic structure and carry a pragmatic message, in other words, the intonation annotation system proposed here integrates information about dynamic and pragmatic features. The intonational annotation system in development for Brazilian Portuguese (henceforth BP) attempts to combine these two pieces of information in the same annotation system, by using tools and methodologies inspired by intonational phonology [5][10][18] intonation workshops [5][19], and by the dynamic systems research paradigm [11].

This work analyzes BP intonational features using intonational annotation, which allows us to observe the relationships between (1) an annotation system which includes dynamical contours, (2) the precise alignment between tones and contours with the phonetic chain.

2. The ToBiPI system

The ToBiPI– *Transcription of Brazilian Portuguese Intonation* – is a system developed for BP intonational annotation [12][13]. This system intends to represent the BP intonation features using a function-oriented dynamic analysis [9][20]. For example, the ToBiPI system dynamic orientation, leave the phrase accents (see also this decision for European Portuguese (henceforth EP) [8]) and the boundary tones

annotation, and allow for the precise alignment between tone and stressed vowel, or between the maximum velocity peak of rising and falling alignment with the stressed vowel. Although we assume that this system applies a dynamic annotation, the uses of H and L and the detection of prominences only by workshops form, such as the ToBI [5][19] system, are aspects of intonational phonology research paradigm [10][18].

Presently, the ToBiPI annotation uses the methodological tools of both research programs [11], however, the application of this annotation allows a functional representation of fundamental frequency (f_0) as it correlates to spatial and temporal sentence organization.

2.1 Pitch accents

The participation of transcribers (researchers and students) of Campinas State University (Unicamp) and São Paulo Catholic University (PUC) participated in 8 workshops that led to the development of the ToBiPI system [4][12][13]. The outcome of these workshops was a set of eight labels for pitch accents annotation, shared on static and dynamic tones, as shown in the following table:

Table 1: *set of tones of ToBiPI system*

Static Tones	
L	low
H	high
Dynamic Tones	
LH	rising
>LH	late rising
HL	falling
>HL	late falling
LHL	rising-falling
HLH	falling-rising

These tones are described as follows [2]:

- rising: have high peak reached at the middle of the stressed vowel;
- late rising: high peak after the stressed syllable;
- falling and late falling: same pattern of the rising tones, replacing the low tone as the reference for alignment.
- rising-falling: presents consecutive high tones where the second tone is always lower than the preceding tone;
- falling-rising: two high tones at adjacent syllables, where the second is lower than the first and sharper than the rising-falling tone.
- low and high: static tones, used to signal leveling into the baseline and the topline respectively.

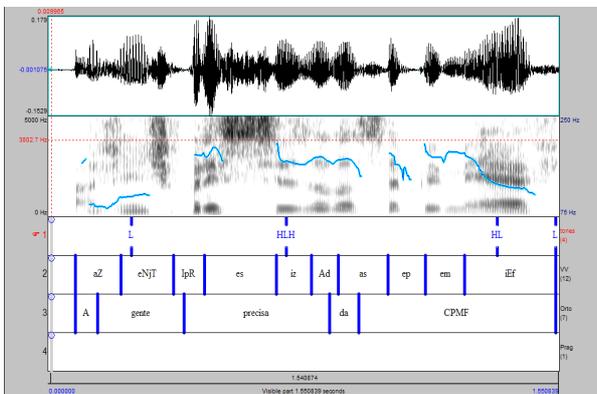
These dynamic tones do not represent an association of two tones, as we can see in some phonological systems, but only one dynamic contour represented by their movement: rising or falling.

The ToBiPI system leaves out the diacritic *, associated with a tone, in favor of an alignment between tones and syllables such as that of a tone target with the CV transition or the maximum/minimum velocity peak [1][21]. This can be automatically tracked from the utterance segmentation into V-toV units [3]. ToBiPI does not use boundary tones. Instead, the system uses same L and H static tones in order to maintain a distinction between form and function. The boundary strength levels are inferred by using a Praat (<http://www.fon.hum.uva.nl/praat>) script, the SGdetector algorithm, that detects duration peaks throughout the utterance [2]. As for the phrase accent, the ToBiPI system consider redundant this annotation, by the fact that in BP the edge tones annotation, with L and H labels, is sufficient to the intermediate and final boundary annotation.

2.2 Annotation Tiers

The manual annotation schema in ToBiPI is composed of 4 tiers in Praat [12][13] and obeys the following order: (i) the first tier is used for the tonal annotation, using the labels described on table 1 and the diacritics ! and | used to signal, respectively, the onset of periods of downstep and upstep; (ii) the second tier is used for sentence segmentation in V-V units; (iii) the third tier is reserved for the orthographic transcription; (iv) the last tier is used for observations about pragmatic information, for example the kind of sentence focus.

Figure 1: example of four levels ToBiPI transcription on Praat. The spontaneous utterance transcribed is "A gente precisa da CPMF" (We need the CPMF)



The last tier, pragmatic annotation, in addition to the dynamic contours, confer a functional feature to the ToBiPI annotation. The Praat window with the ToBiPI transcription layout can be observed on following figure.

3. Speech data

3.1 Spontaneous speech corpus

The corpus used for the intonational annotation is the Speech Prosody Studies Group spontaneous speech corpus,

which is a compilation of radio broadcasts [12]. The corpus consists of close to 8 hours of interviews in mp3 format from *Rádio Você AM* (Americana, Brazil) and *Rádio CBN* (São Paulo, Brazil).

For this paper a set of utterances produced by a male and a female speaker were analyzed. The speakers' professions are economist and psychologist, respectively. Both are in their thirties and are from to the same state in Brazil.

The choice for this corpus was motivated by an international research trend that has been concerned with the analysis of spontaneous speech data [12][16][17]. Spontaneous speech compose a database which we can find several language phenomena in a real situation of use, and its represents an important tool for the phonetic research.

The advantage of spontaneous speech data, rather than lab speech data, is the presence of several different phenomena we can be observed in a natural and informal context. The disadvantage is the impossibility of having sentence repetitions. For this reason, we used different sentences to demonstrate different phenomena, which could be less productive than a comparison between the same sentences.

3.2 Data analysis

For the analysis the mp3 audio files were converted on wav audio files, and segmented on smaller files containing narrow focus. Only the files which present conversations from the interviewees were selected to compose the corpus, the professional announcer speech was discarded.

In the following section, these files were transcribed using the ToBiPI system in the intonation workshops. The employ of an intonational label to describe a specific tone is defined by the transcribers' consensus.

4. Narrow focus in BP

Narrow focus is implemented in BP intonation by *f0* configurations in agreement with spatial and temporal constraints. The general intonational pattern of BP and EP are described by several authors [4][8][14][15], however, most authors present a linear analysis of focus.

The narrow focus features exemplified here are dynamically related to the *f0* alignment and pattern contour, which could be observed by rising and falling tone groups. The falling tones are always preceded by a short rising, and the rising tones by a short falling [2]. These movements seem like an obligatory condition for the realization of rising focus realization in BP. Perceptual tasks confirm this hypothesis by the judgment of several foci whose *f0* height in pre-focal places was manipulated. The judgment shows that the *f0* low positions give us the perception of a narrow focus in comparison of *f0* high positions in the same place. The alignment between *f0* and the linguistic material will determine the local implementation of these patterns. To explain how the dynamic tones are able to represent narrow focus at a temporal and spatial constraint it was selected the following examples of narrow focus annotation in BP.

In the first example we see typical narrow focus where the LH contour starts aligned at the maximum velocity peak which is falling (see the arrow in figure 3) and the stressed vowel is aligned with the high peak (figure 2). This contour is the most common on BP foci. Less frequent than the LH, the >LH tone has the same pattern, but in a late position (figure 4). This late

rising make with f_0 alignment with the stressed vowel don't achieve the high peak, as it is achieved in LH contour. This delay could occur as a consequence of a disalignment between the production and the articulatory system, and produce the perception of a narrow focus.

Figure 2: LH contour aligned with the stressed vowel of "bilhões" at the utterance "quarenta bilhões" (forty billion)

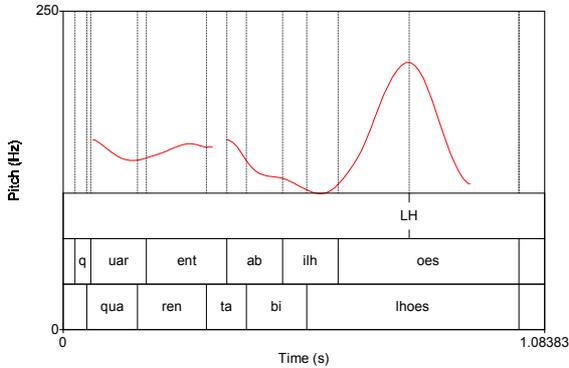


Figure 3: to the maximum velocity peak of falling (F_{max}) and rising (R_{max}) alignment in the same utterance. L and H are, respectively, the high and low peak.

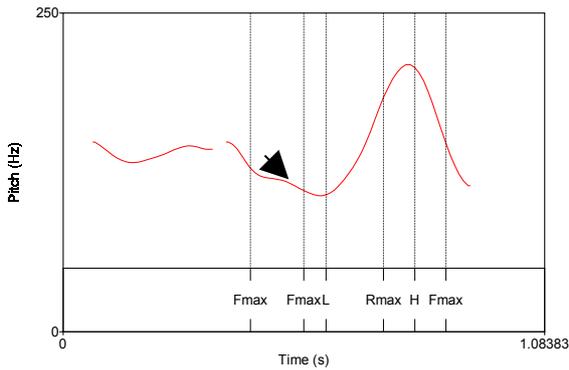
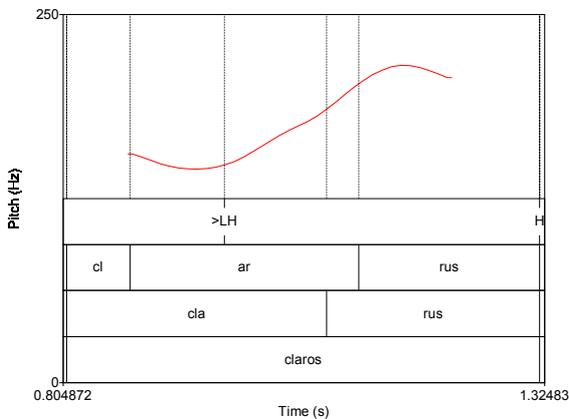


Figure 4: the late rising contour at the word "claros" (clear) where the stressed vowel is aligned with the rising movement.



The label LHL consecutive high tones in a dynamic tone, where the tone aligned with the stressed syllable is lower than

the preceding one, and higher than the following tone (figure 5). The annotation doesn't provide falling among two high tones. The LHL denotes a focus that happens in a low position aligned with stressed syllable, and is generally used to transcribe a weak focus at the final sentence position, in opposition to the HL focus, transcribed an emphatic focus in a stronger fall.

Figure 5: LHL contour example aligned with the stressed vowel of the utterance "chocaram" (they shocked)

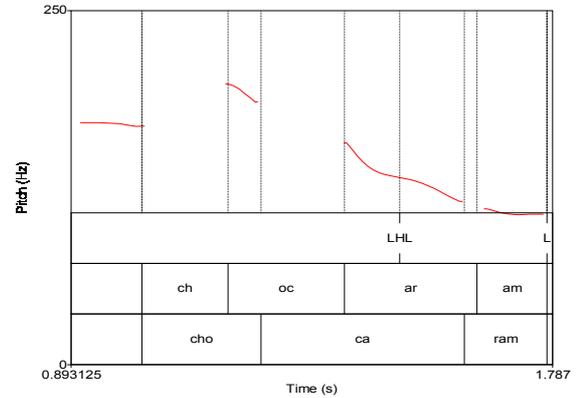


Figure 6: HLH contour aligned with the stressed vowel of "quinto ano seguido" (fifth year in a row).

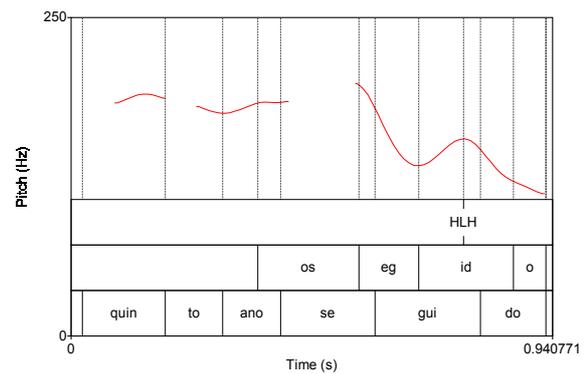
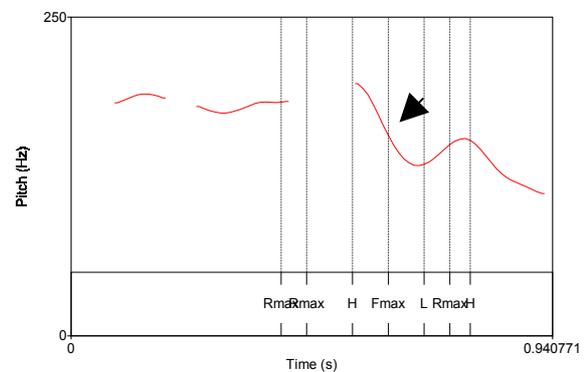


Figure 7: The maximum velocity peak alignment for "quinto ano seguido". Observe the similar alignment at F_{max} as accous on LH.



HLH needs a falling movement between two high peaks to implement the second high tone at the same alignment (figure

6). The first peak is aligned with the secondary stressed syllable, which amplifies the focus effect. One important observation is the alignment implemented by the HLH contour. This is the same that occurs on the LH contour, both are started by the point of maximum velocity peak of falling at the stressed syllable onset (see arrow in figure 7). The main difference between them is that LH is not preceded by the secondary stress that occurs in HLH.

5. Discussion

The existence of functional tonal alignment between the intonational curve and linguistic material controlled by spatial and temporal constraints on narrow focus realizations, show the necessity of a non-linear annotation. To best describe these features is an intonational annotation with a dynamic-functional orientation is necessary, in order to represent processes involved with communicative functions.

For example, the LH rising pattern seems necessary in BP for the realization of narrow focus in sentence medial position. However, it presents some variations in agreement with the specific focus function, as occurs in HLH contour. The same happens with HL falling pattern in final sentence position, which presents LHL contour as variation. For a dynamic intonation annotation the existence of a system that can integrate these features is important for economic labeling that can also describes some functional aspects.

The ToBiPI transcriptions seek to attain these dynamic features. The transcriptions at least seem to point to these results, if we take into account the satisfactory evidence of speaker strategies applied for different foci production, such as alignment and delay, which carry pragmatic-communicative functions [7].

6. Acknowledgements

This work was supported by FAPESP grants 05/58214-0. We thank Meghan Armstrong for the help with the English revision.

7. References

- [1] Atterer, M. and Ladd, D. R., 2004. "On the phonetics and phonology of "segmental anchoring" of /θ/: evidence from German" *Journal of Phonetics* 32, 177-197.
- [2] Barbosa, P.A., 2007. Prominence- and boundary-related acoustic correlations in Brazilian Portuguese read and spontaneous speech. This volume.
- [3] Barbosa, P. A., 2000. "Syllable-timing in Brazilian Portuguese": uma crítica a Roy Major In *D. E. L. T. A.* 16 (2), 369-402.
- [4] Barbosa, P. A. ; Lucente, L. ; Silveira, L. S. ; Madureira, S., 2007. Intonational Patterning in the Paulista Brazilian Portuguese Variety: Two Case Studies in the Framework of AMPER. In: *Actas do III Congresso Internacional de Fonética Experimental*. Santiago de Compostela: Xunta de Galicia, 117-129.
- [5] Beckman, M. E., Hirschberg, J., Pitrelli, J. F., 1994 Evaluation of Prosodic Transcription Labeling Reliability in the ToBI Framework. *Proceedings of the International Conference on Spoken Language Processing*, Yokohama, 18-22.
- [6] Botinis, A., Granström, B., Möbius, B., 2001. Developments and Paradigms in Intonation Research, *Speech Communication* 33, 263-296.
- [7] Dombrowski, E., Niebuhr, O., 2005. Acoustic patterns and communicative functions of phrase-final rises in German: activating and restricting contours. *Phonetica* 62(2-4), 176-195.
- [8] Frota, S., 2000. *Prosody and focus in European Portuguese. Phonological phrasing and intonation*. New York: Garland Publishing.
- [9] Hirst, D. J., 2005. Form and Function in the Representation of Speech Prosody. *Speech Communication* 46, 334-347.
- [10] Ladd, D. R., 1996. *Intonational Phonology*. Cambridge: Cambridge University Press.
- [11] Lakatos, I., 1978. *Falsificação e Metodologia dos Programas de Investigação Científica*. Lisboa: Edições 70.
- [12] Lucente, L.; Barbosa, P., A., 2007. Notação Entoacional do Português Brasileiro em Corpora de fala Semi-Espontânea e Espontânea. *Revista Intercâmbio* 16 (no prelo).
- [13] Lucente, L., Silveira, L. S., Barbosa, P. A., 2006. Declarativas em PB: downstepping ou nova combinação bitonal? H+!H* e H+L*. *IX Congresso Nacional e III Congresso Internacional de Fonética e Fonologia*. Belo Horizonte.
- [14] Moraes, J. A., 1998. Intonation in Brazilian Portuguese, in HIRST, D., DI CRISTO, A. (eds.) *Intonational Systems: a Survey of Twenty Languages*. Cambridge. MIT Press.
- [15] Moraes, J. A., 1999. F0 Declination in Brazilian Portuguese in Read and Spontaneous Speech, *Proceedings of the 14 th International Congress of Phonetic Sciences*, San Francisco, 2323-26.
- [16] Ostendorf, M., Shafran, I., Shattuck-Hufnagel, S., Carmichael, L., Byrne, W., 2001. A prosodically labeled database of spontaneous speech. *Proc. ISCA Tutorial and Research Workshop on Prosody in Speech Recognition and Understanding*, 119-121.
- [17] Peters, B., 2005. The Database: The Kiel Corpus of Spontaneous Speech. In: Klaus J. Kohler, Felicitas Kleber, Benno Peters (ed). *Prosodic Structures in German Spontaneous Speech*. AIPUK 35.a.
- [18] Pierrehumbert, J., 1980. *The phonology and phonetics of English Intonation*. PhD thesis, MIT.
- [19] Silverman, K., Beckman, M., Pitrelli, J., Ostendorf, M., Pierrehumbert, J., Hirschberg, J., Price, P., 1992. TOBI: A Standard Scheme for Labeling Prosody. *Proceedings of the International Conference on Spoken Language* 92, Banff.
- [20] Xu, Y., 2005. Speech melody as articulatorily implemented communicative functions. *Speech Communication* 46, 220-251.
- [21] Xu, Y. and Xu, C. X., 2005. Phonetic realization of focus in English declarative intonation. *Journal of Phonetics* 33, 59-197.