Cues to the perception of prosodic breaks in Brazilian Portuguese

Carolina Serra & Sónia Frota

UFRJ (Brasil) & Universidade de Lisboa (Portugal)

Experimental and Theoretical Advances in Prosody - Cornell University, April 11-13, 2008

Introduction & Background

It has long been known that the location of intonational boundaries is somehow related to the location of syntactic boundaries. The prosodic hierarchy theory captures this relation via syntax-phonology mapping principles which, combined with phonology proper constraints like phrase weight and/or size, predict the preference patterns of intonational phrasing in production and comprehension (Selkirk, 2000; Watson & Gibson, 2005; for Romance languages, Elordieta et al., 2005; Frota et al., 2007). Previous descriptions of intonational breaks within Romance (based on read speech isolated sentence corpora) show that the main boundary cues are a H boundary tone and a F0 rise located either on the stressed or on the boundary syllable (Frota et al., 2007).

The PhP and the IP in Portuguese

Phonological Phrase (PhP) construction: Lexical head + all elements to the left within the Maximal projection of Lex + following XP *complement* containing just one PW (Frota 2000, Tenani 2002). In BP, unlike EP, the PhP is characterized by the regular occurrence of a pitch accent in the most prominent element (Frota & Vigário, 2000; Tenani, 2002).

Intonational Phrase (IP) construction: strings not structurally attached to the sentence tree, as well as any remaining sequence of adjacent PhPs in a root sentence are mapped onto IPs (Frota 2000). Akin to Selkirk's (2005) Comma Phrase. IPs in BP are characterised by the presence of a nuclear contour and a pause (Tenani, 2002). In EP, IPs are characterized by (i) a nuclear accent, (ii) a final boundary tone, (iii) pre-boundary lengthening, (iv) initial tone associated to the right-edge of the first prosodic constituent within de I-phrase and (v) the *loci* of potential pause insertion (Frota 2000, 2003; Vigário 2003).

Present Research

Goals

(1) To capture the relation between prosodic breaks as established by prosodic hierarchy theory (e.g. Nespor & Vogel 1986; for Portuguese, Frota, 2000; Tenani, 2002) and perceived prosodic boundaries in read and spontaneous speech;
(2) To describe the phonetics and phonology of intonational breaks, namely the inventory of tones and the way they are realized (Ladd, 1996);
(3) To assess the most prominent cues to the perception of prosodic breaks in read speech and spontaneous speech.

Main contributions to the field

* New data: Brazilian Portuguese (BP)
* Contribution to the cross-linguistic knowledge about the placement and shape of prosodic boundaries
* Contribution to the understanding of the way prosodic boundaries are realized and perceived
* Contribution to the knowledge about prosodic differences (prosodic phrasing and the various types of boundary cues) across speech styles (e.g. Blaauw 1994)

Method

Corpus: 3 spontaneous speech (SS) and 3 read speech (RS) passages of about 2 minutes each.
Speakers: (1) Production - 3 female speakers. (2) Perception - 11 listeners. University students, born in Rio de Janeiro, aged between 22 and 38.
Procedure:
(1) Production - Interviews in a quiet room at UFRJ's Phonetics Lab. Reading of a passage of the speech uttered during their interview. Passages fully annotated for phonological and intonational phrasing (for Portuguese, Frota 2000, Tenani 2002, Fernandes 2007), to define the placement of **predicted prosodic breaks**. Intonational analysis performed to determine the tonal shape of the nuclear contours, along the lines of the Intonational Phonology approach (e.g. Beckman & Pierrehumbert, 1986; Ladd, 1996; Frota, 2002). PRAAT (Version 4.3.12) has been used.
(2) Perception - Listeners had to signal the prosodic breaks they perceived in each passage. Breaks should be signaled on a piece of paper containing the orthographic transcription of each passage, without any kind of punctuation marks. To assess consistency, listeners repeated the task in 2 sessions (for each session the items were randomized differently). The task was preceded by a training phase. Instructions mentioned that *any kind of break* should be marked.

Measurements:
Location of silent pauses
Lengthening patterns in the nuclear word: duration of stressed and post-stressed syllables relative to the pre-stressed syllable, expressed as percent values.
F0 variation in the nuclear word: F0 measurements obtained at the peak of maximum intensity in the syllable (PMI) were the basis for the computation of F0 variation between stressed and pre-stressed syllables, and F0 variation between post-stressed and stressed syllables, both expressed as percent values.
Type & frequency of nuclear contours (nuclear PA and boundary tone)

Comparison between predicted phrasing and perceived phrasing
Breaks perceived by at least 9 of the 11 listeners > perceived breaks
Properties of predicted AND perceived breaks
Properties of predicted but NOT perceived breaks
Properties of perceived breaks NOT predicted as IPs

Results and Discussion

- **Perceived** prosodic breaks occur at **IP**-boundaries in both styles (Fig. 1).
- More predicted IPs are perceived as breaks in RS than in SS (Table 1).

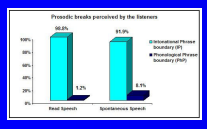


Figure 1.

Table 1 - Distribution of IP-edges perceived and not perceived as breaks

	IP-edges perceived	IP-edges not perceived	All IP-edges predicted
RS	80 (58,4%)	57 (41,6%)	137
SS	57 (41,6%)	80 (58,4%)	137

- Lengthening patterns and F0 properties at the IP-edge – Speaker-dependent strategies (Figures 3-6)

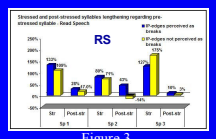


Figure 3.

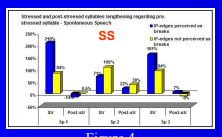


Figure 4.

The bulk of lengthening goes on the stressed syllable (as previously reported for EP in Frota 2000). It **may** be a cue to perception, dependent on the speaker (e.g. Sp1, SS). Overall, bigger lengthening contrasts in **SS**.

Phonetic and Phonological Properties of predicted & perceived / non-perceived breaks

- Presence of pause (Fig. 2).

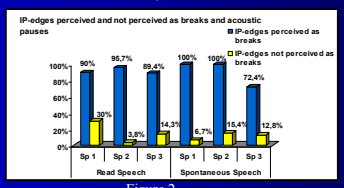


Figure 2.

The **main cue** to the perception of a break is the presence of a pause, across speakers and for both styles

Fig. 5.

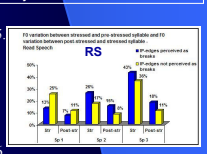
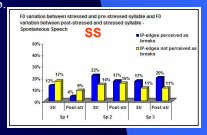


Fig. 6.



F0 variation **may** be a cue to perception, but also dependent on the speaker (e.g. Sp2, RS). Speaker's strategies seem to either favour lengthening or variation (e.g. Sp1 vs. Sp2). Overall, larger F0 variation (on the stressed syllable) in **RS**.

- Nuclear contours – H+L* L% is the most common contour in both styles, but with different frequencies (Fig. 7: **RS** > **SS**)

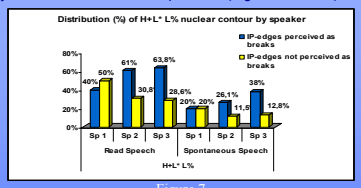


Figure 7.

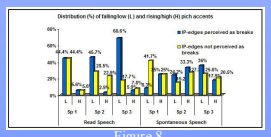


Figure 8.

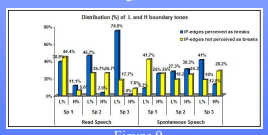


Figure 9.

Falling/low nuclei are dominant in **RS**, as well as the L boundary tone (against previous results for read isolated sentences – Frota et al. 2007); in **SS** the distribution of falling/low and rising/high nuclei, as well as low/high boundaries is about even (Figs 8-10).

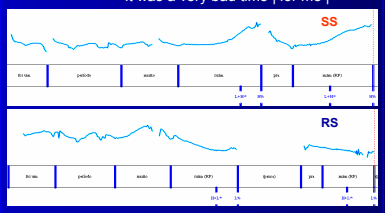
Falling/low nuclei and the L boundary are a feature of perceived breaks in **RS**.

Realization of predicted IPs

Only 14% (38/274) of all predicted IPs were not realized as such (no boundary tone); 20% in **SS**.

Realization of perceived breaks not predicted as IPs: Only 0.8% (6/748) of all PhPs were perceived as breaks (1.3% in **SS**). These were produced as IPs.

Fig. 10. IP-edge with perceived break, Sp 1 – SS / RS
[foi um período muito ruim pra mim] IP
It was a very bad time | for me |



Summary

Perceived breaks occur at IP-edges, regardless of style
More predicted IPs NOT perceived in SS (58% vs. 42%)
Properties of predicted & perceived IPs / not perceived

1. a pause is the strongest cue
2. lengthening (stressed V) is speaker-dependent and characterises especially **SS**
3. F0 variation (stressed V) is speaker-dependent and characterises especially **RS**
4. Nuclear contours: more falling contours in **RS**; more contour variation (token, not type) in **SS**

Conclusion

SS shows more variability in the relation between predicted, perceived and produced phrasing (as expected). The variability emerges from a more even distribution of the same types of cues, that seem to play similar roles in both styles. This even distribution may be an important factor behind the lower perception of predicted breaks in SS.

Unlike in previous reports (e.g. Howell & Kadi-Hanifi 1991, Blauuuw 1994), BP data does not show marked differences in type between styles (such as pauses mainly in SS or at different locations, or a tendency to produce much more phrases in one of the styles relative to the other).

In BP, predicted IP-phrasing is fairly robust across styles: 86% all predicted IPs were realized as such (92% in RS; 80% in SS); only 0.8% of predicted PhPs were perceived as breaks and uttered as IPs (0.3% in RS; 1.3 in SS).

Future work: combined cues & perception

http://www.fl.ul.pt/LaboratorioFonetica/ carolinaserra@letras.ufrj.br sonia.frota@mail.telepac.pt