https://doi.org/10.35520/diadorim.2012.v12n0a3969

ARANTES, Pablo; LIMA, Maria Luiza Cunha; BARBOSA, Plinio A. **Some prosodic correlates of referential status in Brazilian Portuguese**. *Revista Diadorim / Revista de Estudos Linguísticos e Literários do Programa de Pós-Graduação em Letras Vernáculas da Universidade Federal do Rio de Janeiro*. Volume 12, Dezembro 2012. [http://www.revistadiadorim.letras.ufrj.br]

SOME PROSODIC CORRELATES OF REFERENTIAL STATUS IN BRAZILIAN PORTUGUESE

Pablo Arantes (Capes/Prodoc, UFMG)

Maria Luiza Cunha Lima (UFMG)

Plinio A. Barbosa (UNICAMP)

ABSTRACT: In this paper, we present an experiment to investigate if and how referential status is prosodically encoded in Brazilian Portuguese. Brand new referents were compared to given referents along with a control non-coreferential condition in a controlled reading experiment. We also varied word size. We analyzed the following acoustic parameters: fundamental frequency, duration, spectral emphasis and long-term average spectrum. Results show that overall saliency (newness) tends to receive more acoustic prominence in the form of a high initial fundamental frequency rise and longer target word duration. This effect is sensitive to the number of pre-stressed syllables on the target word: the longer the word, the more pronounced the initial F0 rise becomes. Our experiment helps to fill in a gap in Brazilian Portuguese studies and we hope it will contribute to open new research avenues on the theme of prosody-information relationship.
KEYWORDS: Referential Status, Prosody, Brazilian Portuguese

RESUMO: Neste artigo, apresentamos um experimento que investiga se e como o status referencial é expresso prosodicamente em Português Brasileiro. Referentes novos foram comparados a referentes dados, ao lado de uma condição controle com referentes não co-referenciais. Variamos também o tamanho das palavras. Analisamos os seguintes parâmetros acústicos: frequência fundamental, duração, ênfase espectral e espectro médio de longo termo. Os resultados mostram que a saliência referencial (novos) em geral tende a receber maior proeminência na forma de elevação da frequência fundamental e maior duração na palavra alvo. Esse efeito é sensível ao número de sílabas pré-tônicas da palavra alvo: quanto mais longa a palavra mais pronunciada a elevação inicial de F0. Nosso experimento contribui

para preencher uma lacuna da descrição do Português Brasileiro e, esperamos, poderá contribuir para a abertura de novas perspectivas de estudos sobre o tema da relação prosódia-estrutura informacional.
PALAVRAS-CHAVE: Status referencial, prosódia, Português Brasileiro.

Introduction

This article reports the results of an experiment carried out to describe the influence of information structure in a set of acoustic correlates of prosody. The general outline behind the idea of information structure is that the linguistic form of sentences, of which prosody is one component, is affected by hypotheses held by the speaker concerning the "statuses of the mental representations of the referents of linguistic expressions in the mind of the receiver at the moment of utterance" (Lambrecht, 1996, p. 3). Information structure is a very complex subject with a number of interacting aspects influencing how it is actually expressed in sentences. Keeping track of referents in memory – in other words, expressing and following referential status – is one of those aspects and it is, on its own, a complex subject (for a review, c.f. Lambrecht, 1996; Erteschik-Shir, 2007). Choice of referent form, determining systems, choice of syntactic order and verbal voice all play a role in referential status.

Until recently, one of the least studied aspects of information structure was the prosodic marking of referential status, but the subject has received growing interest in the last few years (Baumann & Grice, 2006; Schumacher & Baumann, 2010). In Brazilian Portuguese, despite the existence of previous work on information structure (Stein, 2005; Moraes, 2006) the prosodic expression of referential status is yet to be described. In this work, we aim at starting to fill this gap.

1. Referential status

When interacting, one of subjects' main goals is to exchange information. A sentence can convey new information or presuppose given information. Sentence segments can receive focus because they are either new information or the sentence topic. Besides the overall informational status of propositions within sentences, there is the narrower subject of how known (or given) are the referents in discourse. Information is a concept related to the propositions entertained during an interaction: one must distinguish between new and given information and the entities that can be referred to while expressing those propositions. In other words, there is a distinction between new and given information and given and new referents.

In the present work, we deal with this last specific aspect of the flow of information in discourse: its referents' informational status. In any given discourse, interlocutors have to introduce referents and constantly keep track of them in memory. For each new referent, interlocutors have to assign an address in the discourse model and they may have to get back to this address and update it as discourse unfolds. Language uses different referential forms to signal to a listener/reader where, in memory, to search for the referent, by signaling either that the referent is new and a new address is therefore necessary or that it is already given or at least accessible.

For instance, in the sentence[1] "**A friend of mine** arrived in the city today. *She*'s an architect" the referent "a friend of mine" has not been mentioned before and cannot be recovered from context. This kind of referent is called brand new and is usually introduced by an indefinite article. Recently mentioned referents can be referred back to using semantically lighter expressions such as pronouns or null pronouns as "she" in the example above. In other situations, the discourse model renders previously unmentioned referents accessible, as in the case of associative anaphora in the sentence "**My car** broke down this morning. *The engine* is not working", in which the referent "the engine" is identifiable by association with the previously introduced "car" and therefore preceded by a definite article – that usually accompanies a given referent.

Referents, thus, can have different status ranging from brand new in one extreme to those that are simultaneously given and in focus in the other. Those statuses are not (as we have already seen) a simple given-new binary contrast, but rather form a complex familiarity continuum. Different taxonomies of this continuum have been proposed (e.g. Prince, 1981; Gundel, Hedberg, & Zacharsky, 1993), all of them trying to systematically relate referential forms and referential status, along with different cognitive and discursive models of the underlying reasons for referential form choice and the very existence of this continuum (Almor, 1999; Gordon & Hendrick, 1998; Ariel, 1988). One general finding is that there is an inverse relationship between referent salience in memory and the amount or weight of the linguistic expression used to refer to it: the more salient a referent is the less linguistic effort is put into identifying it.

Referential continua are expressed in different ways, first and foremost through the different forms to encode a referent. Referent forms include pronouns or null pronouns, full NPs determined by different articles, demonstrative pronouns and quantifiers. Prosodic patterns seem to be yet another way to convey the salience of the information. For instance, in West Germanic languages such as English, German and Dutch, research has described a clear relationship between new information and prosodic saliency – higher pitch and often longer duration (Wolters, 1999; Baumann, 2006; Fowler &

1. In the examples given in this paragraph, antecedents are printed in boldface and anaphoric expressions are printed in italics.

Housum, 1987). In contrast, given information is deaccented and shorter. Baumann (2006) proposes that, at least for German, more subtle referential status (such as associative and accessible) receive distinct prosodic patterns as well.

The interaction between prominence and referential status plays a role in discourse processing as well. If prosodic profiles of different referential statuses are swapped, in other words, if new referents are artificially made non-prominent and given referents made prominent, this manipulation disturbs speech processing in a way detectable both in behavioral tasks (Terken & Nooteboom, 1987) and electrophysiological data (Li, Hagoort, & Yang, 2008). Prosodic prominence not only directs attention to a particular word in a sentence, but also interacts in subtle ways with semantic and pragmatic expectations.

To better describe this interaction in BP, we designed a simple, controlled experiment, with the two extreme and most easily identifiable elements of the referential status continuum: the brand new, introduced by an indefinite article, and the given information, introduced by a definite article and referring back to the topic/subject of the previous sentence. This was done to investigate the maximum referential status contrast and to keep our description relatively model-free, since the distinction used here is not problematic within referential status models. After generating a corpus to contrast these two referential statuses, we analyzed different acoustic parameters to determine which ones contribute the most to mark referential status prosodically.

2. Experiment

Some authors tend to analyze information structure through spontaneous or semi-spontaneous speech corpora. Work with this kind of corpora certainly presents advantages, but we agree with Xu (2010) on the powerful role corpora of carefully designed sentences have to uncover the underlying mechanism of speech production. Therefore, we asked our subjects to read texts with controlled linguistic variables.

The experiment corpus consists of carrier sentences in which test-words appear in three conditions regarding their referential status: as new referents, given referents and in a control condition. In all status conditions, test-word position and syntactic function are the same: the nominal phrase (target NP) in which each test-word is embedded is always sentence-initial and the sentence's grammatical subject carrier. In our corpus' sentences, the same noun appears at the beginning of each sentence. Referential status is also marked by the contrast between the indefinite article in the first noun occurrence and the definite article in the second. In (1), the indefinite NP "A journalist" in the first sentence introduces a new referent in the discourse that is referred back to by the definite NP "the journalist" in the second sentence, which is thus a given referent.

(1) [Um jornalista]$_{NEW}$ sentou para escrever sua coluna no último minuto. [O jornalista] $_{GIVEN}$ estava sem ideias naquele dia.
[A journalist]$_{NEW}$ sat down to write his column at the last minute. [The journalist] $_{GIVEN}$ ran out of ideas that day.
Example 2 illustrates the occurrence of the word *jornalista* 'journalist' in the control condition:

(2) [O jornalista] passou por aqui hoje.
[The journalist] came by today.

Strictly speaking, the initial NP in (2) introduces a new referent. However, there is no other referent referring back to it later on, in contrast to what happens in the first example, thus "the journalist" in (2) is not part of a referential chain. Nakatani (1999) classifies new referents into two groups: referents that initiate referential chains (as in example 1) called "first mention referents" and referents that are mentioned only once (example 2) and are labeled "single mention referents". It is in that sense that we call this situation a control condition: we expect to show prosodic encoding differences between referents that are within or outside a referential chain.

Appendix A shows the eighteen target words and appendix B lists all carrier sentences. Target words were controlled for the number of pre-stressed syllables (two, three and four) and the stress type (lexical stress on the last syllable – oxytone words – or second to last syllable – paroxytone words –, which are the two most common word stress patterns in BP). In total, 270 repetitions of carrier sentences were obtained (3 word sizes × 2 stress patterns × 3 target words × 3 referential statuses × 5 repetitions).

In carrier sentences of the type exemplified in (1), the average number of syllables between the end of the new NP and the onset of the given NP is 12.6 (with a standard deviation of 2.25). To test the hypothesis that this number is homogeneous among the three groups of carrier sentences defined by words with the size number of pre-stressed syllables, a one-way analysis of variance was performed and showed no significant difference between the three groups [$F(2, 15) = 0.2$, ns]. Having comparable distance between referents across the corpus is important, because Fowler and Housung (1987) suggest it affects the prominence level of the second occurrence of the referent.

The speaker is a male graduate student at State University of Campinas at age 39 at the time of the recording. He was born and lives in a city 100 miles northwest of São Paulo state's capital city. Test sentences were presented in written form in pseudo-randomized order by means of a slide presentation software and were read aloud by the subject. Besides test items, 36 distractor sentences which contained the test-words in sentences with different structures were also included in each slide deck. All five repetitions of the test sentences were recorded on the same day. The slide presentation was randomized before each of the five repetitions. A ten-minute pause was observed between each repetition session. The recording was made in a silent room at State University of Campinas Linguistics Department. Data acquisition was done through a Digidesign Mbox2 sound card using a Behringer B-2 Pro unidirectional microphone. The speech signal was digitized at a sample rate of 44.1 kHz, bit depth of 16 in mono mode and stored directly in digital format in a portable computer.

3. Acoustic analysis

3.1. Duration

The interval considered for analysis was the duration of target words within the test NPs. The interval corresponding to the determiner was not included because definite and indefinite determiners in BP have a different number of syllables or a different syllable structure (*o, a* for definite and *um*, *uma* for indefinites) and including their duration would break the desired *ceteris paribus* condition. Since a correlation between the number of syllables in a word and its duration is to be expected, speech rate was used as a normalization procedure to minimize the influence of word size and stress[2] on the variable being measured. Following common practice, speech rate was measured in syllables per second as the ratio between the number of syllables in a test-word and its acoustic duration.

Target word duration was measured manually using both waveform and spectral information. Location of word boundaries was stored in accompanying metadata files in a per sound file basis for further processing. A custom Praat script was used to extract the duration of the intervals relative to the test-words.

3.2. Fundamental frequency (F0)

The fundamental frequency contour of each sound file was obtained by means of a script that implements a heuristic which attempts to reduce estimation errors (e.g., octave halfing or doubling) by optimizing the choice of floor and ceiling values passed to Praat's autocorrelation F0 extraction algorithm. Remaining errors were hand corrected. Praat Pitch files corresponding to individual sound files were stored for further processing. Details about the different analysis that were carried out will be given in the ensuing sections.

2. Oxytone words in the corpus are always one syllable shorter when compared to paroxytone words having the same number of pre-stressed syllables.

For all F0 analysis, the interval examined was the whole target NP (determiner plus noun). Location of all syllable boundaries was determined using the same criteria mentioned in the previous section (3.1). Metadata files containing syllable boundaries location were also stored for later use.

3.2.1. Central tendency and dispersion measures

The aim of this analysis is to investigate the possible influence of the different referential status values on the overall features of the F0 contour. Simple measures of central tendency (arithmetic mean) and dispersion (standard deviation and range) were used as estimators of F0 variability. Mean F0 can be interpreted as the typical value used by the speaker in each referential status level. Standard deviation estimates how variable the contour is by taking the average of all deviations above or below the mean F0. Range also measures variation but only takes into consideration the highest and lowest F0 values in a particular interval. Praat's built-in query functions were used to extract mean and standard deviation values (in Hertz) of each Pitch file in the analysis interval. The value of range in semitones was obtained by the formula $12*\log_2(max_{Hz}/min_{Hz})$, where max_{Hz} and min_{Hz} are respectively the maximum and minimum F0 values in Hertz.

In the analysis of central tendency and dispersion, the data of words of both stress types were collapsed because we are not interested in possible differences caused by the stress type difference[3].

3.2.2. Time-normalized F0

F0 contours were time-normalized according to the procedure described in Arantes (2011)[4]. By applying that procedure we obtained a mean normalized contour for each referential status condition out of a number of raw F0 contours each having a different duration. Having mean time-normalized contours for the three status conditions allowed the visual comparison of features like the presence, magnitude and alignment of F0 movements.

The number of time-normalized samples per syllable in the target NP was set to five, resulting in a mean intrasyllabic interval duration of 29 ms (SD 15 ms). Prior to normalization, all contours were smoothed using Praat's built-in smoothing function with bandwidth parameter set to 4 Hz.

Each time-normalized contour comprises a number of F0 samples that is a multiple of the number of F0 values taken at each syllable, set here to five. The total number of samples is determined by the syllable count in the test NPs:

3. In oxytone words, the rising F0 movement aligned to the stressed syllable can spread over the first syllable of the following word. Because of that, mean F0 and standard deviation may be slightly lower if the following syllable is not included in the measurement.
4. The procedure described in Arantes (2011) is a slightly modified version of that adopted by Xu (1993).

- words with 2 pre-stressed syllables
    - paroxytones (5 syllables, 25 samples),
    - oxytones (4 syllables, 20 samples)
- words with 3 pre-stressed syllables
    - paroxytones (6 syllables, 30 samples),
    - oxytones (5 syllables, 25 samples)
- words with 4 pre-stressed syllables
    - paroxytones (7 syllables, 35 samples),
    - oxytones (6 syllables, 30 samples)

For each of the six groups described above, the time-normalized contour of each status level is the mean of 15 different contours: 3 target words × 5 repetitions.

3.3. Spectral measures

F0 is by far the most common acoustic parameter investigated in the literature examining the relation between referential status and prosody. Duration has received some attention (e.g. Fowler & Housung, 1987; Wolters, 1999) but other acoustic correlates of prosodic prominence have been neglected for the most part. In order to fill this gap we looked into two parameters that relate changes in voice source to changes in the speech signal spectrum: spectral emphasis and the long-term average spectrum.

3.3.1. Spectral emphasis

The definition of spectral emphasis we adopted here refers to the relative contribution of the high-frequency band to the overall intensity of the speech signal spectrum. Changes in glottal activity due to increased vocal effort can cause the high-frequency band to contribute to a greater degree to the overall intensity than otherwise. A non-exhaustive list of sources of vocal effort increase includes factors such as:

- non-linguistic: distance between speaker and hearer (Traunmüller & Eriksson, 2000);
- affective/stylistic: the maintaining of increased tension in vocal folds in order to speak in a high F0 level characteristic of some emotions and speaking styles can lead to increased vocal effort;
- linguistic proper: increased vocal effort can be a correlate of linguistic categories such as focus and accent (Sluijter, van Heuven & Pacilly, 1997; Heldner, 2003).

A Praat script described in Arantes (2011) that implements the algorithm described in Traunmüller & Eriksson (2000) with additions suggested in Heldner (2003) was used to calculate spectral emphasis for each vowel in the test-words. Only oral vowels in target words were measured. Nasal diphthongs in target words like "peregrinação" and "sofisticação" were excluded because it is unclear how the algorithm will perform on nasal or nasalized segments. Number of oral vowels in the corpus: [a, ɐ] (312), [e] (176), [i, ɪ] (247), [o] (179) and [u, ʊ] (134).

3.3.2. Long-Term Average Spectrum

The Long-Term Average Spectrum (LTAS) shows average sound intensity level in a range of frequencies and it is thought to reflect both glottal and vocal tract characteristics that are relatively independent from the segmental composition of the speech material being analyzed. Increase in loudness will in general affect the high-frequency band more than the low band, lessening the energy difference between them. Following references cited by Master et al. (2006), we compared energy difference between bands 0-0.5 kHz and 2-4 kHz of the spectra of the three referential statuses to see if an unusual difference could be found. Differences in the number and location of peaks in the LTAS spectrum can also be informative, although significant differences are more common when comparing modal voices to non-modal or unhealthy voices, and for that reason we do not expect them to appear here.

Praat's pitch-corrected LTAS extraction algorithm (Boersma & Kovacic, 2006) was used to obtain a spectrum for each referential status level. For the analysis, sound files corresponding to the duration of test NPs of all repetitions were concatenated in three separate ensemble files, one for each status level. Duration of the ensemble files was above 40 seconds (new: 76.9 s, given: 65.8 s, control: 75.6 s), a value regarded as the threshold above which the influence of segmental factors on the resulting spectrum is considered negligible (Master et al. 2006). Silence and consonants were not removed from the sound files because they do not affect the LTAS in the range analyzed here: 0 to 5 kHz. Default parameters of Praat's LTAS extraction function were used, except for minimum and maximum F0 values that were individually adjusted for the three concatenated status files and bandwidth, which was set to 125 Hz.

4. Statistical analysis

Independent variables are: referential status (status) with three levels (new, given and control), word size measured by the number of pre-stressed syllables (prst) in the test-words (2, 3 and 4 syllables), stress type (stress) of test-words with two levels (paroxytone and oxytone).

Dependent variables were speech rate (syllables per second) in the duration analysis, mean F0 and standard deviation (in Hertz) and range (in semitones) in the F0 analysis and spectral emphasis (in decibels).

The main hypothesis tested in this experiment predicts that new referents will be more prominent than given ones, and control referents will be either in an intermediary position between the two or close to new referents since technically they are new referents that are not referred back to. The specific predictions made by the main hypothesis for each acoustic parameter will be presented in the results section.

Analysis of variance was used to test differences in mean between levels of the independent variables. Tukey's HSD (Honestly Significant Difference) test was used to perform multiple comparisons when the independent variable had more than two levels. Details on the analysis done for each acoustic parameter are presented in the results section. An alpha level of 5% was applied throughout all tests, carried in the R statistical computing environment.

5. Results

5.1. Speech rate

Figure 1 shows target word mean speech rate broken down by factors PRST and STRESS in the three levels of STATUS. In order to assess the influence of the linguistic variables on speech rate, a three--way ANOVA was performed having PRST, STRESS and STATUS as independent variables and speech rate as the dependent variable. If the main hypothesis of more prominent new referents is true, we expect lower speech rate for new referents when compared to given and control either in between or close to new referents. Results show that there are significant main effects for all the independent variables [PRST: $F(2, 418) = 29$ $p < .001$, $\eta_p 2 = .1$; STRESS: $F(1, 418) = 229$, $p < .001$, $\eta_p 2 = .31$; STATUS: $F(4, 418) = 33$ $p < .001$, $\eta_p 2 = .21$]. The main effects of prst and stress suggest that the normalization procedure was not entirely successful in neutralizing the influence of word size and stress on speech rate. The longest target-words (4-prst) were spoken at a slower rate when compared to the shorter ones (4-PRST < 2- and 3-PRST, $p < .001$) as were the oxytones when compared to the paroxytones). Additional one-way ANOVAS having status as independent variable were then carried out separately for each combination of the other two variables, prst and stress. Table 1 presents the results.

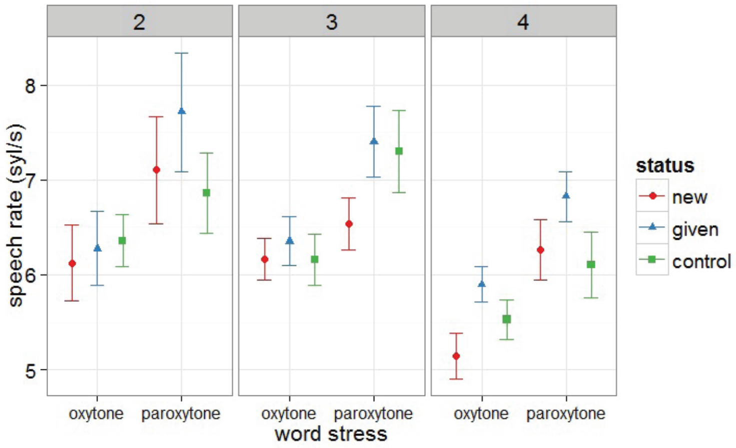

Figure 1: Mean speech rate (syllables per second) broken by word size (vertical panels indicate the number of pre--stressed syllables), word stress and referential status. Whiskers indicate 95% confidence intervals around the mean.

Pre-stressed syllables (PRST)	Stress type (STRESS)	Referent status (STATUS)
2	Oxytone	$F(2, 42) = 0.4$, ns, , $\eta_p^2 = .1$
	Paroxytone	$F(2, 42) = 2.5$, $p < .1$, $\eta_p^2 = .02$
3	Oxytone	$F(2, 42) = 0.7$, ns, $\eta_p^2 = .2$
	Paroxytone	$F(2, 42) = 6.4$, $p < .001$, $\eta_p^2 = .03$ new < control: $p < .05$ given = control: n.s. given > new: $p < .01$
4	Oxytone	$F(2, 42) = 12$, $p < .001$, $\eta_p^2 = .2$ new < control: $p < .05$ given > control: $p < .1$ given > new: $p < .001$
	Paroxytone	$F(2, 41) = 5.8$, $p < .01$, $\eta_p^2 = .4$ new = control: n.s. given < control: $p < .01$ given > new: $p < .05$

Table 1: Speech rate statistical analysis. The third column reports the one-way ANOVA result for each combination of variables PRST (3 levels) and STRESS (2 levels). When a main effect of STRESS was detected, a multiple comparison test (Tukey HSD) was applied to determine statistically significant differences between means of the three levels of STRESS (given, new and control). *P*-values for each comparison are shown.

The new-given contrast followed our hypothesis in three out of six possible PRST and STRESS combinations, the short words (2-PRST and 3-PRST oxytones) failing to show any difference in speech rate between new and given, suggesting that the more pre-stressed syllables a word has, the more reliable the shortening of given referents is.

When the difference between new and given referents is statistically significant, the behavior of the control condition is mixed: control paired with new in the 4-PRST paroxytones, was in between new and given in 4-PRST oxytones and paired with given in the 3-PRST paroxytones, an outcome not predicted by our main hypothesis.

5.2. Fundamental frequency central tendency and variability

This section presents the analysis results of F0 central tendency (5.2.1), standard deviation (5.2.2) and range (5.2.3). Mean values of F0 mean, standard deviation and range for each level of STATUS and PRST variables are shown in Figure 2.

Our main hypothesis predicts that new referents will have higher F0 mean, standard deviation and range values when compared to given referents and that the control condition will present values similar to new referents or will be in an intermediary position between new and given.

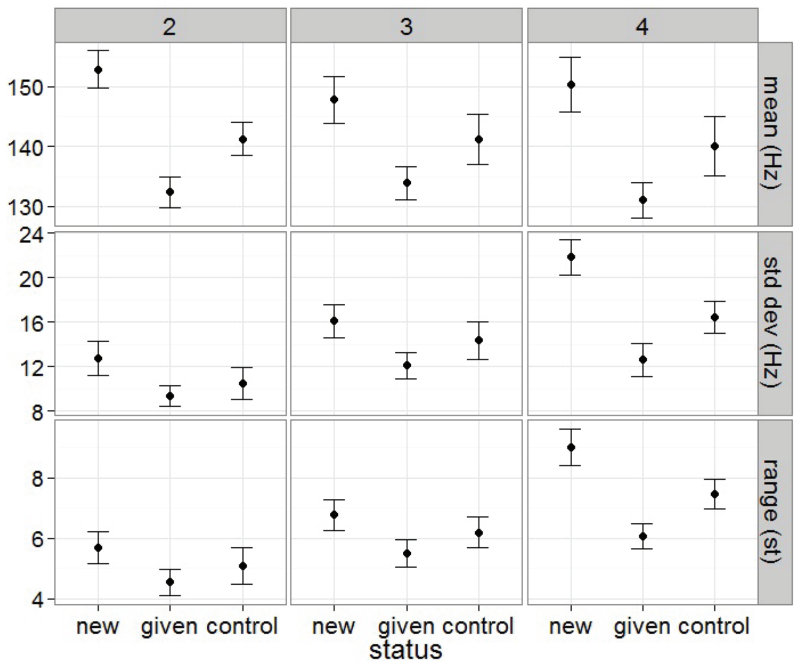

Figure 2: Mean values of F0 mean, standard deviation and range (panel rows), broken by word size (panel columns) and referential status (horizontal axis). Whiskers indicate 95% confidence intervals around the mean.

5.2.1. Central tendency

A two-way ANOVA tested the effect of PRST (all three levels) and STATUS on mean F0 in the test NPs. The results showed a main effect of STATUS [$F(2, 260) = 71.55$, $p < 0.001$, $\eta_p^2 = .35$]; no main effect of PRST [$F(2, 260) = 0.65$, ns, $\eta_p^2 = .005$]; no interaction [$F(4, 260) = 1$, ns, $\eta_p^2 = .02$].

The statistical results confirm what can be seen in the first row of panels in Figure 2: word size does not affect the mean value of F0 in the test NPs; referential status on the other hand has a strong effect on F0 mean. Tukey HSD tests show a statistically significant positive 18 Hz difference between new and given ($p < .001$) referents, a positive 9.5 Hz difference between new and control ($p < .001$) and a negative 8.4 Hz difference between given and control ($p < .001$), confirming our main hypothesis.

5.2.2. Variability: standard deviation

A two-way ANOVA tested the effect of the PRST (all three levels) and STATUS variables on mean F0 standard deviation in test NPs. Results indicate main effects of PRST [$F(2, 260) = 51.7$, $p < .001$, $\eta_p^2 = .29$] and STATUS [$F(2, 260) = 43.14$, $p < .001$, $\eta_p^2 = 0.25$] and significant interaction [$F(4, 260) = 4.9$, $p < .001$, $\eta_p^2 = .07$].

A look at the second row of panels in Figure 2 helps to understand the main effect of PRST. Mean F0 standard deviation tends to go up with the increase in the number of syllables in the test-words, which is confirmed by the multiple comparisons tests: 3-PRST:2-PRST (3.3 Hz, $p < .001$), 4-PRST:2-PRST (6.1 Hz, $p < .001$), 4-PRST:3-PRST (2.8 Hz, $p < .001$. The fact that this tendency does not have a uniform effect on all levels of STATUS (new is more affected) is the source of the STATUS-PRST interaction. Separate ANOVAS for each level of PRST followed by multiple comparisons were employed to compare differences in means among the levels of STATUS (difference in Hz between levels is shown before the *p*-values for each paired comparison):

- 2-PRST: $F(2, 87) = 6.6$, $p < .01$, $\eta_p^2 = .13$; new-given (3.4, $p < .01$), given-control (-1.2, ns), new-control (2.3, $p < .1$)
- 3-PRST: $F(2, 87) = 7$, $p < .01$, $\eta_p^2 = .14$; new-given (3.9, $p < .001$), given-control (-2.3, $p < .1$), new-control (1.8, ns)
- 4-PRST: $F(2, 86) = 36.9$, $p < .001$, $\eta_p^2 = .46$; new-given (9.2, $p < .001$), given-control (-3.8, $p < .01$), new-control (5.4, $p < .001$)

5.2.3. Variability: range

A two-way ANOVA having PRST (all three levels) and STATUS as independent variables and mean F0 range in semitones as dependent variable was used to test the effect of referential status on range. Results indicate main effects of PRST [$F(2, 260) = 63.8$, $p < .001$, $\eta_p^2 = .33$], STATUS [$F(2, 260) = 35.4$, $p < .001$, $\eta_p^2 = .21$] and a significant interaction [$F(4, 260) = 3.8$, $p < .01$, $\eta_p^2 = .05$].

Words of different sizes are not homogeneous. A look at Figure 2 suggests that mean F0 range values tend to grow higher with the increase in the number of syllables in the test-words, which is confirmed by the multiple comparison tests: 3-2 (1 st, p < .001), 4-2 (2.4 st, p < .001), 4-3 (1.4 st, p < .001). Multiple comparisons were employed to compare differences in means among the levels of the STATUS variable separately for each level of that variable (differences in semitones between levels are shown before the *p*-values for the paired comparisons):

- 2- PRST: given-control (-.5 st, ns), new-control (.6 st, ns), new-given (1.1st, p < .01)
- 3- PRST: given-control (-.7 st, ns), new-control (.6 st, ns), new-given (1.3st, p < .01)
- 4- PRST: given-control (-1.4, p < .01), new-control (1.6 st, p < .001), new-given (3 st, p < .001)

5.3. Time-normalized F0 contours

Figure 3 shows time normalized F0 contours for all levels of STRESS (paroxytone and oxytone words), PRST (test-words with 2, 3 and 4 pre-stressed syllables) and STATUS (new, given and control). Contours correspond to the entire test NPs, meaning it includes the determiner and the test-word. Five values were sampled at regular intervals in each syllable. No statistical analysis was performed on the data. In the following paragraphs a visual interpretation of the contours is provided.

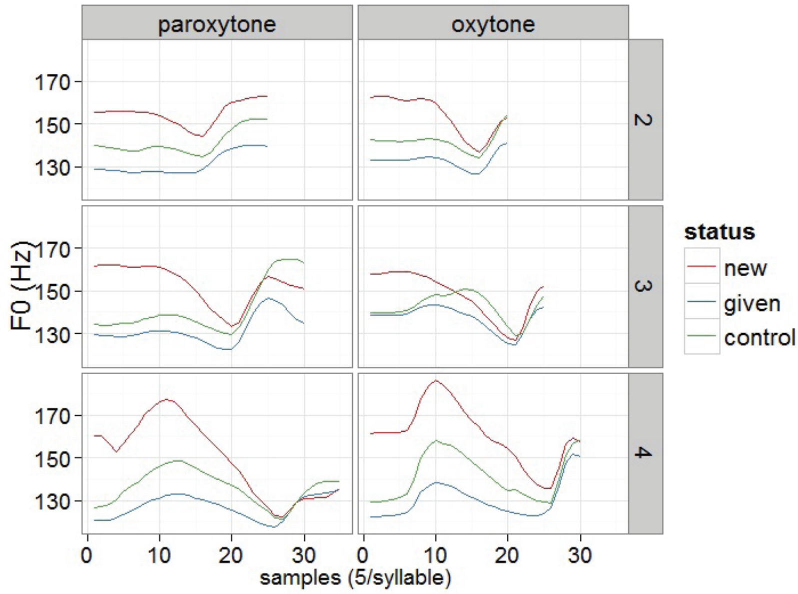

Figure 3: Mean time-normalized F0 contours (values in Hz) of the test NPs broken by stress type (columns) and size of the pre-stressed syllable chain (rows). Line color indicates different referential status. F0 samples are displayed in the horizontal axis - each five samples correspond to a syllable in the test NP.

One common feature in all contours is a rising (final) movement that starts roughly at the beginning of the stressed syllable. In paroxytone words the movement will usually reach its peak at the post-stressed syllable. In oxytone words, because F0 was not measured beyond the stressed syllable, it is not always possible to see the movement reach its peak, which usually will occur in the following syllable. There is no sizable difference among the rising movement aligned to the stressed syllable in terms of excursion when comparing new, given and control referents. The exceptions are the 2-PRST and 3-PRST (paroxytone words) groups, in which the movement in given referents begins at a lower point and do not rise to the same height reached in new and control referents.

The major differences between contours of different status levels lay in what happens at the beginning of the test NP. New referents' contours start in a high plateau (2- and 3-PRST) or rising towards a peak (4-PRST) that is then followed by a falling movement that reaches a minimum at the start of the stressed syllable, where the rising movement described in the previous paragraph starts. A gradient lowering of the initial plateau level or peak height can be observed in the contours of control and given referents.

Word size interacts with referential status in how contours are shaped. The significant factor seems to be that speakers have to fulfill two contradictory goals when producing a contour for a new referent: achieve a high pitch target early in the test NP and a low pitch target prior to the stressed syllable. The low pitch target is followed by a second high target/plateau aligned to the end of the stressed syllable and ensures F0 is low enough for a rise that will stretch over the stressed syllable (and post-stressed when it exists).

When there is a larger number of pre-stressed syllables available for a falling movement, as is the case with the 4-PRST group, the initial pitch target for new referents can be higher than the targets for control and given conditions. Because there is more room for F0 movements, the target is reached through a peak movement instead of a plateau. The low target level seems to be defined by the height of the previous peak. Comparing the contours for new referents in the 4-PRST word group makes it possible to see that in the paroxytone words the initial peak is just below 180 Hz and the valley after the fall is near 120 Hz. In the oxytone words, the initial peak is almost 190 Hz and the low valley is over 130 Hz. The whole contour spans almost 60 Hz or 7 semitones relative to the lowest value.

Compare this pattern to what happens in the 2-PRST group. In the new contours, the initial plateau extends through the two initial syllables and there is only one syllable left for the falling movement to reach the low target that precedes the stressed syllable. The range covered is within 20 Hz, one third the range observed in the longest words in the corpus. The way out of this time constraint is to move the whole new contour up: the ranges of the new and given contours do not overlap for the paroxytones and barely touch in the oxytones.

In general it seems possible to say that the three levels of the STATUS variable elicit distinct contours that differ mainly in the height of the initial pitch target. New contours have the highest target, given the lowest and the control condition sits in the middle. The final shape is affected by the number of intervening syllables between the high target and the following low target. F0 movements aligned to the stressed syllable are not important to the referential status distinction. The differences observed can be seen as a result of the interaction between the pitch targets and the number of pre-stressed syllables.

5.4. Spectral measures

5.4.1. Spectral emphasis

A two-way analysis of variance test having PRST and STATUS as independent variables and spectral emphasis (in dB) as dependent variable was performed. There is a main effect of PRST [F(2, 1034) = 18, $p < .001$, $\eta_p^2 = .03$], no main effect of STATUS [F(2, 1039) = 1.6, ns, $\eta_p^2 = .003$] and non-significant interaction [F(4, 1039) = 0.04, ns, $\eta_p^2 = .0002$].

The lack of a STATUS main effect suggests that, at least in the way it was measured here, spectral emphasis plays no significant role as an acoustic correlate of referential status. The presence of a small-sized but significant effect of PRST indicates that between levels of that variable there must be at least one significant difference. For this analysis, spectral emphasis values for each vowel in the test-words of each level of variable PRST were pooled together. In a second analysis, three separate one-way ANOVA tests were carried out for each level of the PRST variable having the position of the vowel in the test--word (POSITION) as the independent variable (2-PRST: 4 positions, 3-PRST: 5 positions and 4-PRST: 6 positions).

Figure 4 shows mean spectral emphasis by position for the three levels of variable PRST. It is possible to notice a syntagmatic effect, i.e., a change in emphasis value influenced by the position of the vowel in the word. Vowels in stressed syllables have higher mean emphasis than those in pre- and post-stressed syllables, hinting that it can be a correlate of primary word stress[5]. Multiple comparisons results are listed below (the lexically stressed syllable position is printed in boldface):

- PRST-2: [$F(3, 308) = 16$, $p < .001$, $\eta_p^2 = .14$], 1=2=4<**3**, all $p < .001$
- PRST-3: [$F(4, 324) = 28$, $p < .001$, $\eta_p^2 = .25$], 5<1=2=3<**4**, all $p < .001$
- PRST-4: [$F(5, 401) = 11$, $p < .001$, $\eta_p^2 = .12$], 1=2=3=4=6<**5**, all $p < .001$

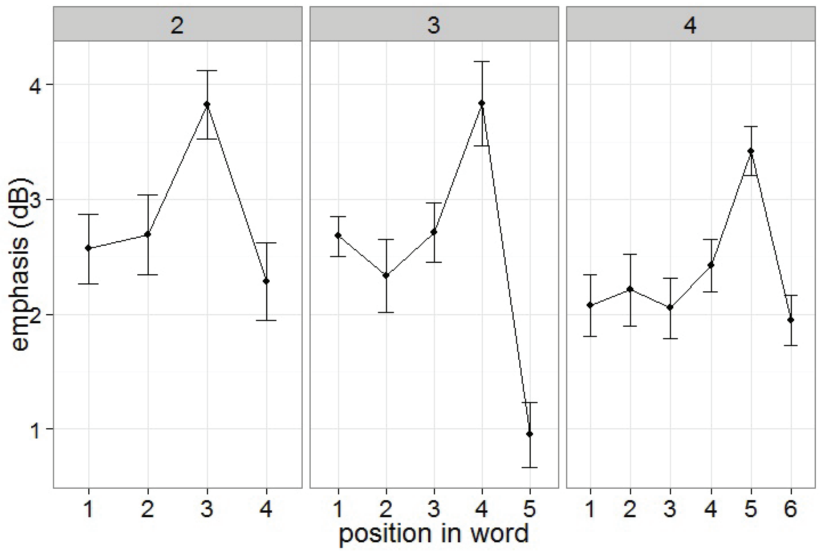

Figure 4: The vertical axis shows mean spectral emphasis, measured as the contribution (in dB) of the high-frequency band to the overall intensity of vowels. The horizontal axis shows the syntagmatic position of vowels in the test-words of different sizes, classified by the number of pre-stressed syllables (2, 3 and 4). Paroxytone and oxytone words are pooled together. Whiskers indicate 95% confidence intervals around the mean.

5. Barbosa et al. (in press) analyzed a host of acoustic parameters that speakers of BP use to cue stress contrasts and reported that spectral emphasis seems to be systematically used as a correlate of word stress across different speaking styles.

5.4.2. Long-term Average Spectrum

Figure 5 presents the LTAS spectra of new, given and control referents in the experiment corpus. The three spectra are greatly similar, in both the number and location of peaks and energy levels. Inspecting the spectra makes it possible to see roughly one peak every thousand Hz, beginning at 1500 Hz. It is common to find one peak in the 0-1 kHz band that is absent in this speaker. The spectrum of new referents has one extra small peak around 2 kHz but the perceptual significance of this peak is not straightforwardly determinable. The overall energy distribution along the frequency bands is also similar and typical of speech, the low-frequency band having higher levels of energy when compared to the high-frequency end of the spectra. Energy differences between 0-0.5 kHz and 2-4 kHz for the three spectra are not striking: -27.4 dB (new), -26 dB (given) and -26.4 dB (control). Based on the results it is possible to say that voice quality differences as they are usually revealed by the LTAS do not seem to play a strong role in cueing referential status in Brazilian Portuguese.

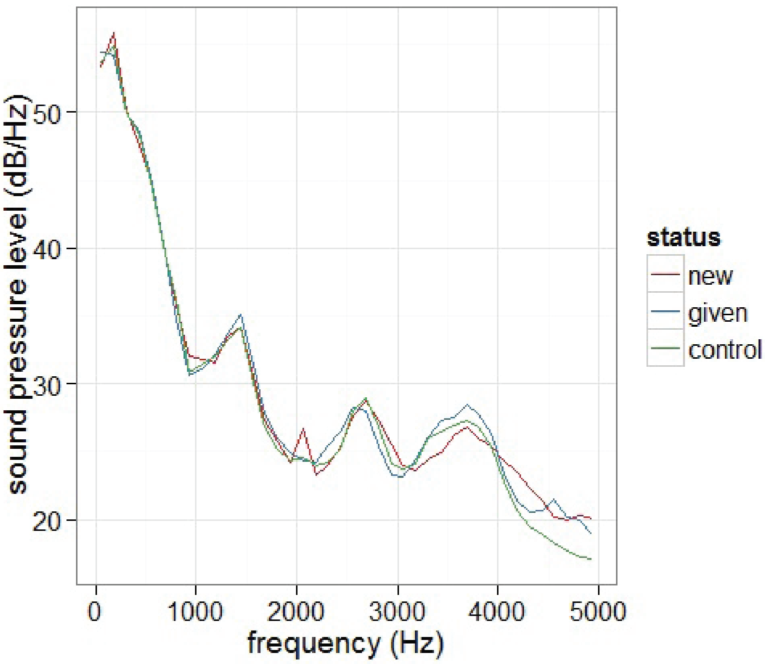

Figure 5: The vertical axis shows power spectral density (dB/HZ) and the horizontal axis shows frequency (Hz). Spectra of the three referential statuses are indicated by line type and color.

6. Discussion

The results presented in this study point to an important role played by duration and fundamental frequency as acoustic correlates of referential status in Brazilian Portuguese, replicating results found in the literature for languages like English, German and Dutch, among others. Confirming our main hypothesis, new referents have a higher degree of salience, cued mainly by slower speech rate, raised F0 level and broader F0 variation around the mean level. As discussed in the results section, word size seems to affect on word duration and F0, suggesting that in future studies this variable should be included.

Our study is one of the first controlled descriptive accounts of how prosody acts as correlate to referential status in Brazilian Portuguese. Most studies on the relationship between referential status and prosody tend to focus heavily on F0 as the main acoustic parameter studied. We investigated a broader range of parameters including duration, spectral emphasis and voice quality features. This wider selection was useful to rule out some parameters that were not sensitive to the phenomenon (spectral emphasis and long-term average spectrum) but more importantly, it revealed duration as a relevant dimension that should not be overlooked in future studies. Moreover, our negative results also contribute to a more informed choice of parameters to be considered.

The methodology we presented can be used as a framework to study the prosodic expression of referential status beyond the initial picture outlined here. In fact, we are currently engaged in expanding the current analysis by exploring new levels of complexity, specifically how the interaction between referential status and syntax affects prosody.

However, we acknowledge limitations to the scope of the reported results. We restricted the current account to production but a natural development will be to investigate if listeners are sensitive to the differences found in acoustics. In addition, our corpus consisted of read speech. It would be important to study other speech genres, such as dialogues, although we believe, following Xu (2010), that more controlled experiments with read speech are important starting points to the description of any linguistic phenomena in speech. Another possible shortcoming is the fact that we only recorded one speaker and therefore just one Brazilian Portuguese variety. We have already addressed this concern in a follow-up study in which we recorded four speakers of another BP variety and our present findings were replicated (Delfino, Cunha Lima, & Arantes, 2012).

We conclude from the results of the experiment reported here that in Brazilian Portuguese informational salience in the given-new distinction is systematically encoded in speech by prosodic means, specifically fundamental frequency and duration.

7. Acknowledgments

This research was supported by Fapesp grant 06/50166-9 to the first author and CAPES grant AUX-PE-PRODOC 2552/2010 to the Language Studies Graduate Program at Universidade Federal de Minas Gerais.

References

Almor, A. (1999). Noun-phrase anaphora and focus: The informational load hypothesis. *Psychological Review, 106*, 748-765.

Arantes, P. (2011). Implementação em Praat de algoritmos para descrição de correlatos acústicos da prosódia da fala. *Anais da II Jornada de Descição do Português* (pp. 32-38). Cuiabá-MT: Sociedade Brasileira de Computação.

Ariel, M. (1988). Referring and accessibility. *Journal of Linguistics, 24(1)*, 65-87.

Barbosa, P., Eriksson, A., & Akesson, J. (in press). Cross-linguistic similarities and differences of lexical stress realization in Swedish and Brazilian Portuguese. *Proc. of XI Nordic Prosody Conference*, August 2012, Tartu, Finnland.

Baummann, S. (2006). The intonation of givenness – Evidence from German. Linguistische Arbeiten, 508. Tübingen: Niemeyer.

Baummann, S., & Grice, M. (2006). The Intonation of Accessibility. *Journal of Pragmatics, 38(10)*, 1636-1657.

Boersma, P., & Kovacic, G. (2006). Spectral characteristics of three styles of Croatian folk singing. *Journal of the Acoustical Socienty of America*, 1805-1816.

Delfino, A. D., Cunha Lima, M. L., & Arantes, P. (2012). Prosodic Marking of referential status in Brazilian Portuguese: A Preliminary Study. *GSCP 2012.*

Erteschik-Shir, N. (2007). *Information Structure: The Syntax-Discourse Interface.* Oxford: Oxford University Press.

Fowler, C., & Housum, J. (1987). Talker's signaling of "new" and "old" words in speech and listeners' perception and use of the distinction. *Journal of Memory and Language*, 489-504.

Gordon, P. C., & Hendrick, R. (1998). The representation and processing of coreference in discourse. *Cognitive Science, 22*, 389-424.

Gundel, J., Hedberg, N., & Zacharsky, R. (1993). Cognitive Status and the Form of Referring Expressions in Discourse. *Language, 69*, 274-307.

Heldner, M. (2003). On the reliability of overall intensity and spectral emphasis as acoustic correlates of focal accents in Swedish. *Journal of Phonetics*, 39-62.

Lambrecht, K. (1996). *Information Structure and Sentence Form: Topic, Focus, and the Mental Representations of Discourse Referents* . Oxford: Oxford University Press.

Li, X., Hagoort, P., & Yang, Y. (2008). Event-related potential evidence on the influence of accentuation in spoken discourse comprehension in Chinese. *Journal of Cognitive Neuroscience*, 906-915.

Master, S., De Biase, N., Pedrosa, V., & Chiari, B. M. (2006). O espectro médio de longo termo na pesquisa e na clínica fonoaudiológica. *Pró-Fono Revista de Atualização Científica*, 111-120.

Moraes, J. A. (2006). Variações em torno de tema e rema. *Cadernos do IX Congresso Nacional de Lingüística e Filologia*, (pp. 279-289). Rio de Janeiro.

Nakatani, C. H. (1999). Prominence variation beyond given/new. *Eurospeech' 99*, (pp. 547-550).

Prince, E. F. (1981). Toward a taxonomy of given-new information. In: P. Cole, *Radical Pragmatics* (pp. 223-255). New York: Academic Press.

Schumacher, B. P., & Baumann, S. (2010). Pitch accent type affects the N400 during referential processing. *Neuroreport, 21(9)*, 618-622.

Sluijter, A. M., van Heuven, V. J., & Pacilly, J. J. (1997). Spectral balance as a cue in the perception of linguistic stress. *Journal of the Acoustical Society of America*, 503-513.

Stein, C. C. (2005). Aspectos da manifestação prosódica em relação com os processos de ativação cognitiva no português. *Ciências & Cognição*, 2-18.

Terken, J., & Nooteboom, S. G. (1987). Opposite effects of accentuation and deaccentuation on verification latencies for given and new information. *Language and Cognitive Processes*, 145-163.

Traunmüller, H., & Eriksson, A. (2000). Acoustic effects of variation in vocal effort by men, women, and children. *Journal of the Acoustical Society of America*, 3438-3451.

Xu, Y. (1993). *Contextual tonal variation in Mandarin Chinese.* The University of Connecticut: PhD dissertation.

Xu, Y. (2010). In defense of lab speech. *Journal of Phonetics*, 10, 329-336.

Wolters, M. (1999). Prosodic correlates of referent status. Proceedings of the XVIth International Congress of Phonetic Sciences.

Appendix A: Test words

Pre-stressed syllables (PRST)	Stress pattern (STRESS)	Frequency of use (occurrences per million)
2	paroxytone	delegado (97.2)
		deputado (266.8)
		jornalista (81.3)
	oxytone	jogador (239.6)
		militar (191)
		senador (174.3)
3	paroxytone	americano (299.9)
		departamento (109.9)
		secretaria (170.3)
	oxytone	consumidor (126.8)
		governador (271.3)
		procurador (59.8)
4	paroxytone	oportunidade (56.6)
		possibilidade (124.5)
		universidade (215.1)
	oxytone	peregrinação (3.5)
		privatização (63.9)
		sofisticação (6.5)

Table 2: Target words list broken by the number of pre-stressed syllbles and stress type. Word frequencies in a 36-million occurrences corpus of spoken and written Brazilian Portuguese.

Appendix B: Test sentences

The first item in each pair of sentences corresponds to the control condition. In the second item, the first occurrence of the test-word is on the new condition and the second occurrence represents the given condition.

- **deputado**

O deputado passou por aqui hoje.

Um deputado foi depor no plenário da Câmara Federal. O deputado foi evasivo em suas declarações.

- **jornalista**

O jornalista passou por aqui hoje.

Um jornalista sentou para escrever sua coluna no último minuto. O jornalista estava sem ideias naquele dia.

- **delegado**

O delegado passou por aqui hoje.

Um delegado comandou uma operação arriscada de busca na fronteira. O delegado prendeu uma quadrilha perigosa.

- **jogador**

O jogador passou por aqui hoje.

Um jogador foi vendido para um time estrangeiro por muito dinheiro. O jogador deixou saudades na torcida.

- **senador**

O senador passou por aqui hoje.

Um senador foi à tribuna e fez um discurso inflamado. O senador tinha uma opinião forte.

- **militar**

O militar passou por aqui hoje.

Um militar foi achado na selva depois de pedir socorro. O militar ficou perdido na mata por dias.

- **americano**

Os americanos passaram por aqui hoje.

Uns americanos fecharam um negócio lucrativo no Brasil. Os americanos vão plantar cana-de-açúcar em São Paulo.

- **departamento**

O departamento fechou mais cedo hoje.

Um departamento será criado em breve pelo prefeito. O departamento deve cuidar dos assuntos financeiros.

- **secretaria**

A secretaria fechou mais cedo hoje.

Uma secretaria será criada em breve pelo prefeito. A secretaria deve cuidar de assuntos financeiros.

- **governador**

O governador passou por aqui hoje.

Um governador sofreu críticas duras de toda a imprensa. O governador botou parentes em cargos públicos.

- **consumidor**

O consumidor passou por aqui hoje.

Um consumidor fez uma queixa contra uma loja que atrasou a entrega. O consumidor tinha se sentido lesado pela loja.

- **procurador**

O procurador passou por aqui hoje.

Um procurador foi ameaçado pelos criminosos. O procurador fez denúncias graves.

- **universidade**

As universidades fazem prova hoje.

Umas universidades marcaram uma reunião com o ministro. As universidades queriam discutir um aumento da verba.

- **oportunidade**

As oportunidades somem no país hoje.

Umas oportunidades surgiram na agência de empregos. As oportunidades apareceram por conta do Natal.

- **possibilidade**

A possibilidade nasce a cada dia.

Uma possibilidade apareceu de repente na vida de José. A possibilidade partiu da indicação de um amigo.

- **privatização**

A privatização cresce no país hoje.

Uma privatização causou medo entre os empregados. A privatização resultaria em possíveis demissões.

- **sofisticação**

A sofisticação passou na moda atual.

Uma sofisticação surpreendeu a plateia no desfile. A sofisticação parecia ter acabado na moda atual.

- **peregrinação**

A peregrinação saiu daqui cedo.

Uma peregrinação reuniu muitas pessoas em volta da igreja. A peregrinação durou o dia todo.