THE DIALECTAL PROSODIC VARIATION IN BRAZILIAN PORTUGUESE SPOKEN IN BAIÃO CITY (PA):

AN INTONATIONAL ANALYSIS OF DECLARATIVE AND YES/NO INTERROGATIVE CLAUSES FROM AMPER-POR CORPUS

Rosinele Lemos e Lemos[1] & Regina Célia Fernandes Cruz[2]

Secretaria de Educação do Estado do Pará[1] / Universidade Federal do Pará[2]

Rosinele_lemos@terra.com.br / regina@ufpa.br

Abstract

This paper, in line with AMPER (*Atlas Multimédia Prosodique de l'Espace Roman*), aims to present an acoustic analysis of the modal intonation role in Brazilian Portuguese spoken in Baião. Sample data has been drawn from the work of (Lemos 2015) and methodological procedures were grounded in AMPER-POR coordination directives. Speech samples from four native speakers, two males and two females, aged above thirty years old, two with low and two with high school levels were analyzed. The *corpus* is constituted by twenty-one sentences from AMPER-POR *corpus* with both simple and complex nominal phrases at the final position, in both intonational modalities – declarative and yes/no questions. The three Portuguese lexical stresses have been considered on our acoustic analysis of fundamental frequency (F0). The results show that F0 is a strong physical parameter to discriminate the analyzed intonational modalities.

Keywords

AMPER Project, dialectology, Brazilian Portuguese

A VARIAÇÃO PROSÓDICA DIALETAL DO PORTUGUÊS BRASILEIRO FALADO NA CIDADE DE BAIÃO (PA): UMA ANÁLISE ENTOACIONAL DE SENTENÇAS DECLARATIVAS E INTERROGATIVAS TOTAIS DO CORPUS AMPER-POR

Resumo

O presente artigo divulga os resultados da descrição prosódica vinculada ao projeto AMPER relativa à análise acústica do papel da entoação modal no português falado em Baião (PA). Foram aplicados todos os procedimentos metodológicos determinados pela coordenação geral do projeto AMPER-POR na seleção dos informantes, formação do *corpus* e tratamento dos dados. Selecionaram-se amostras de fala de dois locutores femininos e dois masculinos moradores nativos da localidade alvo, com faixa etária acima de trinta anos e de diferentes níveis de escolaridade (fundamental e médio). Analisaram-se vinte e uma sentenças do *corpus* AMPER-POR contendo sintagmas nominais finais simples e estendidos, nas modalidades declarativa neutra e interrogativa total. Os resultados apontam que a F0 é determinante na distinção da entoação modal, com as principais variações de F0 ocorrendo na última sílaba tônica do sintagma nominal final.

Palavras chave

projeto AMPER, dialetologia, Amazônia Paraense

1. Introduction

The dialectal description of linguistic variation at the prosodic level in Brazilian northern region constitutes an object of study and investigation within the scope of the *Atlas Multimédia Prosodique de l'Espace Roman*[1] international project (henceforth AMPER), carried out by the AMPER-North[2] team in charge of providing contributions for the elaboration of the Multimedia Atlas of the Brazilian northern region under the AMPER project.

The AMPER-North team has collected data from nine different municipalities in the state of Pará,[3] namely: a) Abaetetuba (Remédios 2013); b) Belém (Brito 2014); c) Mosqueiro Island (Guimarães 2013); d) Bragança (Castilho 2009); e) Cametá (Santo 2011); f) Curralinho (Freitas 2013); g) Mocajuba (Costa 2015); h) Santarém (Lima, in

---

[1] <http://w3.u-grenoble3.fr/dialeto/AMPER/AMPERfr_fichiers/frame.htm>
[2] There is an institutional project at the UFPA (Federal University of the Pará) named AMPER-North project (Prosodic Multimedia Atlas of Northern Brazilian Portuguese).
[3] Pará state is located in the northern region of Brazil.

progress); and i) Baião (Lemos 2015), from where data has been drawn for the present study.

This paper refers to the findings of (Lemos 2015) regarding the acoustic analysis of the Fundamental Frequency (F0) role – observed to be the most relevant parameter – in distinguishing the modal intonation of the Portuguese language spoken in Baião (PA).[4] For this study in particular, we decided to demonstrate a specific dialectal prosodic variation of declarative and interrogative modalities. We observed intonation patterns of phrasal modalities, taking into account the different types of phrases (AP, NP and PP), lexical stress (proparoxytone, paroxytone and oxytone), schooling social variables (elementary and high school levels), and gender (male and female). Our main goal was to verify the performance of the physical parameter of F0 in distinguishing the target modalities investigated by the AMPER project. For this reason, we observed the F0 performance, relating it to the previously mentioned social and linguistic variables present in the AMPER *corpus*.

To demonstrate the F0 role in distinguishing the phrasal modalities in Baião (PA), we selected from (Lemos 2015) some speech samples of four native speakers, two males and two females, aged above 30 and with different schooling levels (elementary and high school).

To foster better understanding of the performed analysis, this work has been structured as follows: in the introductory section, we present an overview of our research; in section 2, we show the source of the analyzed data (Lemos 2015); in section 3, we describe the methodology adopted for data analysis; in section 4, we present the analysis emphasizing the strong correlation between stress and variation of F0 and in distinguishing intonational modalities of Brazilian Portuguese spoken in Baião (PA); in section 5, we present our final remarks and in section 6, our references.

2. The Data Source (Lemos 2015)

---

[4] Baião is a city located on the northern of Brazil, 204 km from Belém, the capital city of Pará State. Situated by the Tocantins River, Baião is part of the Northeast region of Pará and its access can be done by car or bus through PA-151 highway or by boat through the Tocantins River.

As previously referred, the data used in the present paper were selected from (Lemos 2015), who produced the first prosodic description of Brazilian Portuguese (henceforth BP) variation spoken in the city of Baião, located in the north of Brazil, within the Amazon region. The relevance of the data is related to historical, social, cultural and geographic aspects of Baião, regarded in this paper with the purpose of building the speakers' profiles. Following, we present a brief overview of the research developed by Lemos (2015).

2.1 On the description by Lemos (2015)

In reference to Lemos (2015), data collection was performed in compliance with the criteria established by the AMPER project coordination in terms of speakers' selection, *corpus* building and data treatment. The selected speakers were three native men and three native women, aged between 30 and 75, with different schooling levels (elementary, high school and college). A total of 66 sentences were recorded from AMPER-POR *corpus*, involving the target modalities investigated by the project. Lemos (2015), in her Master's thesis constituted a *corpus* of 6 hours, 30 minutes and 8 seconds of recorded material detailed in Table 1, below:

Speaker	Sex	Age	Schooling Level	Pitch Scale	Record size
BF91	F	75	Elementary school	130 Hz to 500 Hz	1h 23min 39sec
BF92	M	32	Elementary school	90 Hz to 400 Hz	1h 53min 36sec
BF93	F	30	High school	130 Hz to 500 Hz	48min 58sec
BF94	M	48	High school	90 Hz to 400 Hz	53min 50sec
BF95	F	42	College	130 Hz to 500 Hz	39min 29 sec
BF96	M	40	College	90 Hz to 400 Hz	33min 36sec

Table 1. Corpus description (Lemos 2015)

Data treatment methodological procedures included six phases: a) standardized codification of six recorded repetitions from each speaker; b) isolation of these repetitions in individual files; c) segmentation of audio signals in PRAAT software;[5] d) application of *script amper praat* in order to obtain the acoustic measures of vocalic segment of each repetition automatically; e) selection of the three best repetitions; f) application of MatLab[6] interface to obtain the mean of the acoustic parameters of the three best repetitions of each sentence and its modalities.

Acoustic parameters, which were taken into account during the acoustic analysis itself, were the same as those controlled by the AMPER project coordination, which are: F0, duration and intensity. Lemos (2015) used the mean of the acoustic parameters of the three best repetitions of each sentence and target modality, comprised in 0.txt files generated by the MatLab interface. A pitch scale for each speaker was established. As shown on table 1 above, female speakers had a scale pitch established between 130 Hz and 500 Hz, whereas male speakers were between 90 Hz and 400 Hz. F0 values on Hz were converted into semitones (ST) to ensure the comparison of different individuals' samples. Duration was calculated in milliseconds (ms) and intensity in decibels (dB).

According to the outcomes observed by Lemos (2015), F0 is a determinant factor in the process of distinguishing intonational modalities if we consider the Portuguese linguistic variation in Baião. Another finding concerns the role of duration measures, which proved to complement the F0 in differentiating the two analyzed phrasal modalities. On the other hand, intensity did not behave as a strong acoustic factor to distinguish declarative and yes/no questions.

Considering that Lemos' work has already been reviewed, we can, hereupon explain the methodological procedures adopted in this paper.

---

[5] Available in http://www.fon.hum.uva.nl/praat/.

[6] <http://www.mathworks.com/products/matlab/?requested*Domain=www.mathworks.com*>

3. Methodology

The research carried out in Baião, followed the same criteria adopted by (Lemos 2015) explained in the preceding section in what concerns speakers' selection and data treatment.

In order to build up the *corpus* of this paper, we selected 21 out of the 66 sentences of the AMPER-POR *corpus* in both target intonational modalities – declarative and yes/no questions. We analyzed data from four speakers, two males (BF92 and BF94) and two females (BF91 and BF93), whose profiles are described in Table 1. As the target of the acoustic analysis consists in the last word of the final phrase of the sentence, and as each selected sentence contains a representative stress in the nucleus of its phrase, we decided to describe them in table 2 below.

In more specific terms, data have been recorded in the speaker-informants' houses where they were asked to repeat the 42 sentences (21 declarative and 21 interrogative) from the *corpus,* six times each, randomly with no interruption until the last one. A mono and headed tape recorder was used to capture the audio and a notebook was used to project the images in a PowerPoint program.

In reference to data treatment, sentences were labeled in order to identify each speaker. For instance, a sentence labeled BF91 bwta1[7] identifies a female speaker at fundamental school level who speaks Brazilian Portuguese from Pará – Baião and the sentence is identified as declarative in first repetition [*O pássaro nadador gosta do Renato* 'The swimming bird likes Renato']. After codifying the sentences, files were isolated by individuals and the three best repetitions were selected.

For the instrumental analysis procedures, two computational programs were used: (1) Matlab – as it has specific applications developed by Antonio Romano for that purpose, and; Praat – using a "script" and an interface developed by Albert Rilliard.[8]

Considering the TXT files, we elaborated the graphics which show the movements of F0 acoustic parameters in the declarative and interrogative modalities investigated in Baião.

---

[7] B- Brazilian Portuguese; F- Pará state; 9- Baião; 1- Female speaker at fundamental school level.
[8] Script developed especially for the project AMPER by Albert Rilliard, available at:
<https://groupeaa.limsi.fr/membres:rilliard:outils_amper> (Accessed 08 April 2016).

L'une des hypothèses de départ sur lesquelles se base notre démarche est que les variétés linguistiques que nous nous proposons d'étudier présentent (1) des phénomènes d'accentuation, de mise en relief d'unités complexes dans la chaîne segmentale, et (2) des phénomènes de structuration intonative de nature syntaxique, sémantique, pragmatique et expressive. Ayant observé que ces faits se manifestent grâce à l'évolution d'un choix restreint de paramètres acoustiques de durée, hauteur, intensité et, dans une moindre mesure, timbre, on s'est consacrés pendant longtemps à définir des techniques d'extraction des valeurs assumées surtout par les trois premières de ces variables (Romano et al. 2005: 1).

We aimed to describe the dialectal prosodic variation of declarative sentences and yes/no questions of Brazilian Portuguese spoken in Baião by observing the intonational pattern of phrasal modalities considering the adjectival, nominal and prepositioned phrases, lexical stress (i.e. proparoxytones, paroxytones, and oxytones), as well as social aforementioned variables – gender (male and female) and schooling level (elementary and high school). Four speakers at the following ages: BF91 (75 years old); BF92 (32 years old); BF93 (40 years old); and BF94 (48 years old) contributed for the collection of data.

Altogether, 504 utterances were analyzed and systematized as follows: 21 sentences x 3 best repetitions x 2 intonational modalities x 4 speakers.

	CV.cv.cv	cv.**CV**.cv	cv.cv.**CV**
1	O pássaro gosta do **pássaro**. 'The bird likes the bird'	O Renato gosta do **Renato**. 'Renato likes Renato'	O bisavô gosta do **bisavô**. 'The great-grandfather likes the great-grandfather'
2	O Renato gosta do **pássaro**. 'Renato likes the bird'	O bisavô gosta do **Renato** 'The great-grandfather likes Renato'	O Renato gosta do **bisavô**. 'Renato likes the great-grandfather'
3	O bisavô gosta do **pássaro**. 'The great-grandfather likes the bird'	O pássaro gosta **do Renato**. 'The bird likes Renato'	O pássaro gosta do **bisavô**. 'The bird likes the great-grandfather'

4	O bisavô gosta do pássaro **bêbado**. 'The great-grandfather likes the drunk bird'	O pássaro gosta do bisavô **pateta**. 'The bird likes the goofy great-grandfather'	O Renato gosta do bisavô **nadador**. 'Renato likes the swimming great-grandfather'
5	O pássaro gosta do Renato **bêbado**. 'The bird likes drunk Renato'	O pássaro gosta do Renato **pateta**. 'The bird likes goofy Renato'	O pássaro gosta do Renato **nadador**. 'The bird likes swimming Renato'
6	O pássaro gosta do bisavô **bêbado**. 'The bird likes the drunk great-grandfather'	O Renato gosta do pássaro **pateta**. 'Renato likes the goofy bird'	O Renato gosta do pássaro **nadador**. 'Renato likes the swimming bird'
7	O pássaro gosta do Renato **de Mônaco**. 'The bird likes Renato from Monaco'	O pássaro gosta do Renato **de Veneza**. 'The bird likes Renato from Venice'	O pássaro gosta do Renato **de Salvador**. 'The bird likes Renato from Salvador'

Table 2. Twenty one sentences analyzed, demonstrated by lexical stress

The acoustic measures of F0 (Hz) from the 0.txt files containing the mean of the physical parameters of the three best repetitions generated by the AMPER interface were primarily used. The F0 values in Hz were then converted into semitones (ST)[9] by using Albert Rilliard's scale created for the AMPER project. We observed the variations of F0 in the last word of the final nominal phrase of the sentence. Thus, the type of phrase – nominal (NP), adjectival (AP) and prepositioned (PP) – is considered a linguistic variable, in consonance with the type of lexical stress in the analysis here undertaken, that is – proparoxytones (**CV**.cv.cv), paroxytones (cv.**CV**.cv), and oxytones (cv.cv.**CV**). Besides the type of phrases and stress pattern, we also took into account social variables: gender – male (M) and female (F) – and schooling level – Elementary (EF) and High school (EM) – to describe the target modalities.

The values related to F0 (ST) in the word that occupies the last position of the final nominal phrase were organized in an Excel spreadsheet in order to obtain the F0 curves, as described in the next section.

---

[9] We used an Excel page created by Albert Rilliard for the AMPER project members.

4. F0 variation patterns of declarative and total yes/no question modalities identified in the data from Baião (PA)

In this section, we present the results of this research. It is noteworthy that the curves analyzed take into account all the linguistic and social variables aforementioned.

4.1 Curve pattern of F0 for proparoxytone stress pattern

In relation to the proparoxytone stress pattern, we observed three distinct intonational patterns on the final position for the yes/no questions modality and one pattern for the declarative modality. When contrasting the schooling levels and the types of phrases, we observed a significant fall in the F0 movement generally from the mid posttonic syllable, keeping downward until the final posttonic syllable of the sentences, both on declarative and yes/no questions, as observed in Figures 1 and 2, below.
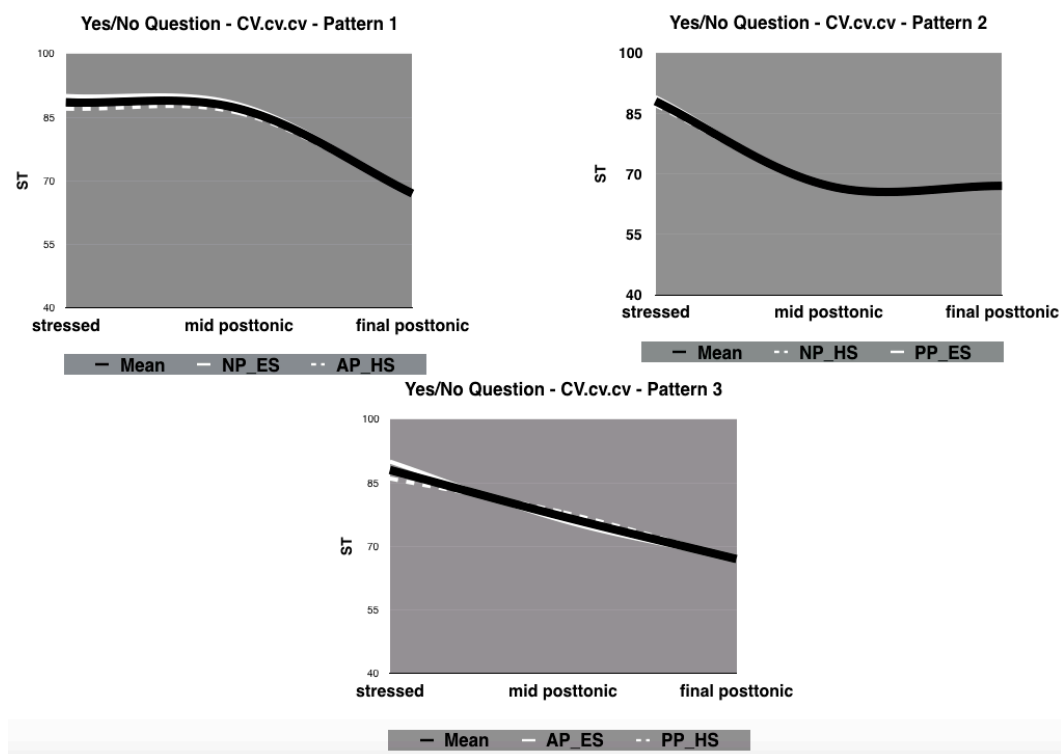

Figure 1. F0 (ST) Patterns for proparoxytone lexical stress of yes/no question. The white lines identify individual F0 curves of each linguistic – NP, AP and PP – and social variable – ES (elementary school level) and HS (high school) – and the black lines show the mean of the F0 variations.

Figure 1 allows the identification of three patterns of F0 curves for yes/no question sentences with proparoxytone word in the final position. The data show that nominal phrases (NP) from speakers with elementary school (ES) level present the same F0 curve pattern as the adjectival phrases (AP) from speakers with high school (HS) level (Pattern 1). Meanwhile, nominal phrases (NP) from speakers with high school (HS) level and prepositioned phrases (PP) from speakers with high school (HS) level present a different pattern (Pattern 2). The only difference between patterns 1 and 2 refers to the F0 movement of mid posttonic syllable of proparoxytone word that occupied the last position of the phrase.

On one hand, mid posttonic syllable in pattern 1 keeps an upward F0 movement that starts in the stressed syllable of the word, whereas mid posttonic syllable in pattern 2 shows a downward movement of the F0 curve. Pattern 3, corresponding to adjectival phrases (AP) from speakers with elementary school level (ES) and prepositioned phrases (PP) from speakers with high school (HS) level, shows a completely different movement compared to the two previous patterns as it presents a constant fall of the F0 starting in the stressed syllable of the word.
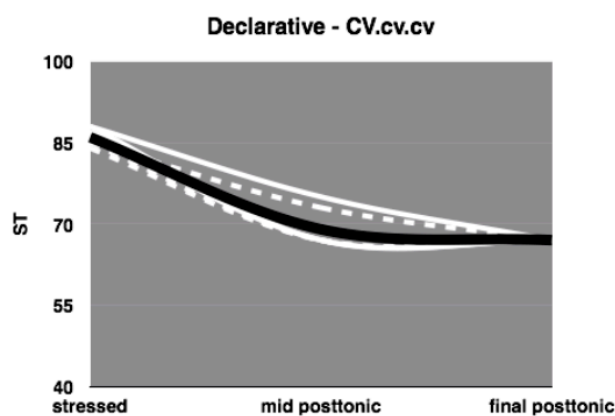

Figure 2. F0 (ST) Patterns for proparoxytone lexical stress of declarative sentences. The white lines identify the individual F0 curves of each linguistic variable - NP, AP and PP – and the social variable - ES (elementary school) and HS (high school) - and the black lines show the mean of the F0 variations.

Although three patterns of F0 movement for yes/no question sentences with proparoxytones words on the final position of the final word were identified, only one

104

pattern was identified for declarative sentences. This single pattern presents downward curves from the stressed syllable of the target word (Figure 2).

The other two Portuguese language stress patterns – paroxytone and oxytone – register only one movement of F0 curve by modality, as observed in the following subsections.

4.2 Curve pattern of F0 for paroxytone stress pattern

Each modality of the sentences containing paroxytone words in the final position of the sentence presented only one pattern per modality, as shown in Figure 3, below.
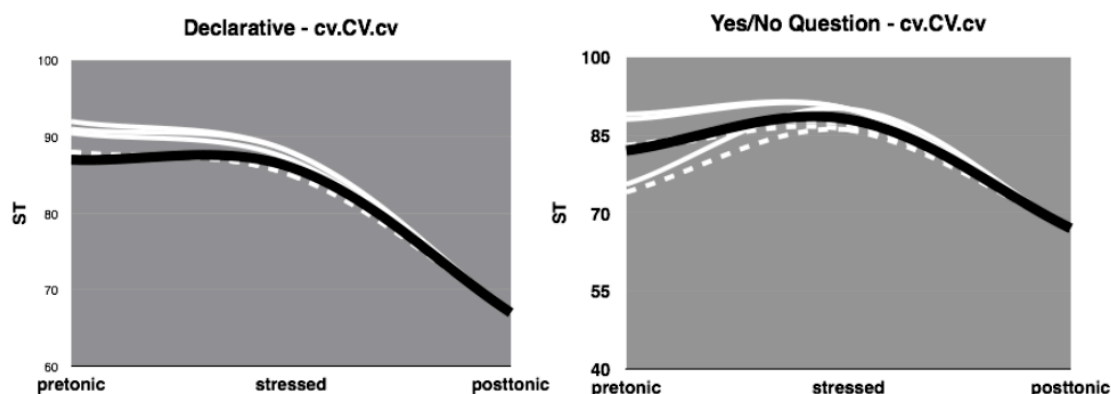

Figure 3. F0 (ST) Patterns for paroxytone lexical stress of declarative sentences and yes/no questions. The white lines identify individual F0 curves of each linguistic and social variables and the black lines show the mean of the F0 variations.

The most important distinctive movements between the two modalities lie either on the stressed syllable or on the final posttonic syllable. Both syllables present the same type of F0 movements, that is, an upward movement on the stressed syllable and a downward movement on the posttonic syllable. The difference, however, rests on the amplitude of the movement. The upward movement of the stressed syllable of yes/no question modality is much higher than that of the stressed syllable, while the downward movement of the final posttonic syllable of declarative sentences is much lower than that of the interrogative. Therefore, we can attest that yes/no question sentences with paroxytone word on the final position present a circumflex pattern of F0 curve.

Although it was possible to establish a distinctive pattern of the phrasal modalities of the sentences with proparoxytone and paroxytone words on the final position, significant results related to the oxytone stress pattern were not verified, as detailed in the next subsection.

4.3 Curve pattern of F0 for oxytone stress pattern

The analysis that took into account the oxytone stress pattern showed a very close intonational pattern for the two target modalities. In both cases, an upward movement of the curve towards the stressed syllable that occupies the final word position was registered, as Figure 4 shows below:
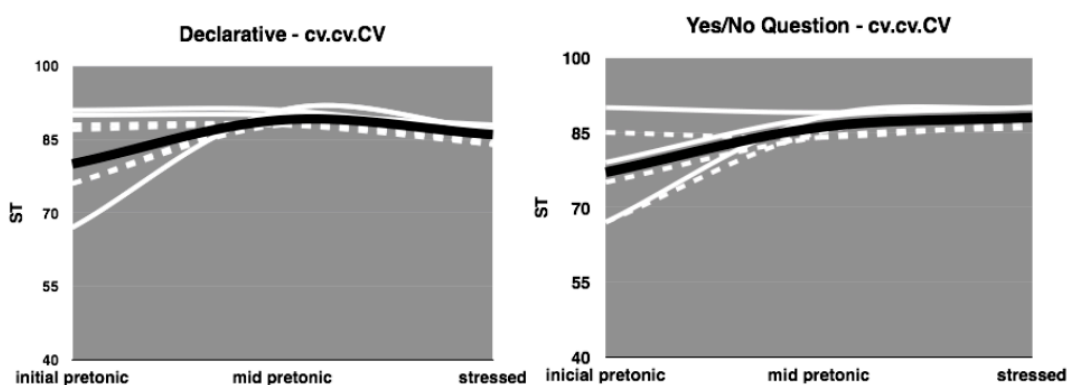

Figure 4. F0 (ST) Patterns for oxytone lexical stress of declarative sentences and yes/no questions. The white lines identify individual F0 curves of each linguistic and social variables and the black lines show the mean of the F0 variations.

The only slight difference observed in relation to a possible F0 movement that distinguishes both modalities is a higher movement on the final part of the stressed syllable of the yes/no question modality.

5. Final considerations

The present study found that F0 is determinant to distinguish the modalities analyzed in this paper, since data revealed that the main F0 variations occur right on the stressed syllables of the final phrases nucleus.

106

According to the analysis here undertaken, eight patterns of F0 curve were identified, one for each target modality containing a sort of lexical stress on the final phrase position. Only the yes/no question sentences with proparoxytone words showed three distinctive patterns. The only difference that might be observed in relation to a possible F0 movement that distinguishes both modalities would be a higher movement on the final part of the stressed syllable on yes/no question modality. Broadly speaking, the descriptions from the Portuguese language carried out by the AMPER team in the North of Brazil have confirmed the main hypothesis posed by the project. That is to say, the most meaningful variations of the acoustic parameter of F0 occur in the last stressed syllable of the nominal phrases in declarative and interrogative clauses according to the studies developed by Remédios (2013), Brito (2014), Guimarães (2013), Castilho (2009), Santo (2011), Freitas Neto (2013), Lemos (2015) and Costa (2015).

The results brought up by Lemos (2015), while studying the variation from Baião, were similar to those found in the AMPER-North literature as they attest declarative sentences present a downward movement in the intonational nucleus of the nominal phrase at the end of sentences, whereas interrogative clauses present an upward curve in that same nuclear region – the gripping movement. Restating, in the scope of AMPER studies within the Amazon portion of Pará state, and regarding investigated variations, F0 shows ascent and descent curves for declarative and interrogative clauses, respectively.

References

BRITO, C. (2014) *Acento versus entoação no português falado em Belém (PA). Análise prosódica dos dados de Belém do projeto AMPER-Norte*, Unpublished Bachelor's degree monography, Belém (Pará): Federal University of Pará.

CASTILHO, F. C. (2009) *Formação de Corpora para o Atlas Dialetal Prosódico Multimídia do Norte do Brasil: Variedade Linguística de Bragança (PA),* Unpublished Bachelor's degree monography, Bragança (Pará): Federal University of Pará.

Cᴏsᴛᴀ, M.S. da S. (2015) *Análise acústica da relação acento versus entoação no português falado em Mocajuba (PA).* Unpublished Master's degree dissertation, Belém (Pará): Federal University of Pará.

Fʀᴇɪᴛᴀs Nᴇᴛᴏ, J. (2013) *Contribuições para a formação do Atlas Prosódico Multimédia do português do Norte do Brasil – AMPER – POR: variedade linguística do Município da Ilha de Marajó (PA).* Unpublished Master's degree dissertation, Belém (Pará): Federal University of Pará.

Gᴜɪᴍᴀʀãᴇs, E. (2013) *A Variação Entoacional da Ilha de Mosqueiro: contribuições para a formação do Atlas Prosódico Multimídia do Português do Norte do Brasil. (Amper – Norte).* Unpublished Master's degree dissertation, Belém (Pará): Federal University of Pará.

Lᴇᴍᴏs, R. (2015) *A Variação Prosódica em Sentenças Declarativas e Interrogativas do Português Falado em Baião (PA),* Unpublished Master's degree dissertation, Belém (Pará): Federal University of Pará.

Lɪᴍᴀ, L. (in progress) *Contribuições para o Atlas Prosódico Multimédia do Português do Norte do Brasil: AMPER – POR: Variedade linguística do município de Santarém (PA),* Unpublished Master's degree dissertation, Belém (Pará): Federal University of Pará.

Rᴇᴍéᴅɪᴏs, I. (2013) *Contribuições para a formação do Atlas Prosódico Multimédia do português do Norte do Brasil – AMPER – POR: variedade linguística do Município de Abaetetuba (PA),* Unpublished Master's degree dissertation, Belém (Pará): Federal University of Pará.

Rᴏᴍᴀɴᴏ, A., J.P. Lᴀɪ & S. Rᴏᴜʟʟᴇᴛ (2005) "La méthodologie AMPER", *Géolinguistique*, hors série 3, 1-5.

Sᴀɴᴛᴏ, I. (2011) *Atlas Prosódico Multimédia do Município de Cametá (PA),* Unpublished Master's degree dissertation, Belém (Pará): Federal University of Pará.

108