PITCH PATTERNS IN BRAZILIAN PORTUGUESE: AN ACOUSTIC-PHONETIC ANALYSIS

Sandra Madureira
Department of Linguistics
Pontifícia Universidade Católica de São Paulo

ABSTRACT- Fundamental frequency contours of declarative, interrogative and imperative sentences are taken into account in describing the prosodic structure of sentences in Brazilian Portuguese.

INTRODUCTION

This paper discusses intonation in Brazilian Portuguese. It is part of an ongoing research on prosody within the project Text-To-Speech System for Brazilian Portuguese. This project which interfaces Linguistics and Engineering represents cooperative work between a group of researchers from the Language Studies Institute and two groups of researchers from the Electrical Engineering Department at the University of Campinas, São Paulo, and its basic aims are to provide the phonetic-acoustic analysis to feed the synthesis and recognition modules and to construct an integrated text and speech signal processing system. The phonetic-acoustic analysis presupposes a phonological theory directed towards the phonetic implementation. This theory being developed by Albano (1992, 1993) and called The Act and Effect Phonology analyses phonological and phonetic events by means of common categories and provides an abstract representation of linguistically or stylistically overruled speech events which can be objectively related to descriptions of the acoustic signal.

The basic aim of this paper is to analyse some features of the fundamental frequency contours of declarative, interrogative, and imperative sentences in Brazilian Portuguese. Fundamental frequency, duration and amplitude are the acoustic features corresponding to the phonetic feature of pitch, but fundamental frequency is its major determinant. Intonation, the pattern of pitch changes over time, has grammatical, textual and emotional functions, the latter playing a more important role as Bolinger (1989) points out and for its description syntactic, semantic, pragmatic and rhythmic analysis are essential. Changes in pitch may cause changes in the quality of voice (Bolinger, 1989; Klatt & Klatt, 1990; Pittam, 1987) which are also relevant for the study of intonation.

The knowledge derived from the analysis of the variation of the fundamental frequency and other prosodic features of the sound segments in the sentences is meant to be activated in the derivation of realistic contours in a text-to-speech system for Brazilian Portuguese so as to improve the inteligibility and naturalness in the synthesized speech which is usually less redundant than natural speech and certainly benefits from prosodic implementation which introduces perceptually relevant cues (Quené & Kager, 1992).

MATERIAL AND METHOD

This analysis takes into account sentences recorded by two male speakers. The sentences which constitute the data corpus were so built up as to include varied number of syntactic constituents, varied number of expansions (modifiers to the heads) within each constituent, varied number of syllables and phonemes. The lexical fields used were those of banking, airport and communication systems mainly.

The data corpus has been processed in a top-down hierarchy: from sentences to segments. All constituents (higher and lower) have been classified into categories. Global as well as local attributes were described. The relative position and the class of the word showing the highest value in a type of sentence were also described.

The sentences were recorded in a studio. For the acoustic analysis of the sentences, the sonograph Kay, model 5000 was used. For the extraction of pitch, besides the sonograph, a software developed in the Electrical Engineering Department at the University of Campinas (UNICAMP) was used. The speech data was digitized to 16 bits/sample at a sampling rate of 16 KHz.

Fo measures were taken at the beginning and end of the utterances and immediately before and after the nonterminal pauses. Local measures were taken at the segmental level in three parts of the syllabic nucleus: beginning, maximum value and end. Based on these values, mean values (Aubergé,1992) have been calculated for the syllable. Duration of the stressed syllables of the content words in the sentence were also measured. Nonterminal pauses were measured and classified.

A prosodic data base has been built up. The prosodic data base comprises the following information: written text, phonetic transcription, pitch graphics, duration and initial and final pitch values of segments. Its design was guided by Larreur et al. (1990) and Aubergé (1992).

RESULTS

A preliminary analysis has shown the characteristic global features of the patterns and the adjustments of F_0 values towards the end of sentences. The findings can be summarized as follows:

- In a simple neutral declarative sentence the maximum value of F_0 is located around the stressed syllable (usually the longest syllable in the sentence) of the content word which is the head of the noun phrase, whenever this happens to have three or more syllables;
- In a simple neutral declarative sentence the maximum value of F_0 is located around the stressed syllable of the verb following the noun phrase subject whenever this happens to have less than three syllables;
- In an imperative sentence the maximum value of F_0 is located on the stressed syllable of the verb in the imperative form;
- The pitch at the beginning of imperative sentences tends to be higher than in declarative sentences;

- In a simple interrogative sentence without question-words (total question) the maximum value of F_0 is located on the stressed syllable of the last content word;
- In a simple interrogative sentence with question-words (partial question) the maximum value of F_0 is located on the question-word;
- In complex declarative sentences the F_0 contour resets at major syntactic boundaries;
- In complex interrogative sentences the F_0 is recurrent;
- In complex sentences where the first clause is imperative the steep pattern will not affect the embedded clause;
- Before nonterminal pauses in a sentence an increase in F_0 value will indicate continuative intonation pattern and a decrease will indicate terminative intonation;
- The decrease in F_0 before nonterminal pauses can be accompanied by breathy voice or laryngealization;
- The decrease in F_0 at the end of utterance is usually accompanied by creaky voice;
- F_0 values usually increase in intensifying adverbs;
- The distance between the peak and the valleys in the last word in an utterance will vary according to the number of syllables: the greater the number of syllables, the greater the range between peaks and valleys;
- The number of syllables tends to be correlated with the value of the initial pitch in the sentence: the greater the number of syllables the highest the initial pitch value in the sentence.

SUMMING UP

As in a large number of languages, the declarative sentences in Portuguese show a declining F_0 global contour and the partial interrogative sentences a rising F_0 global contour. The latter contrasts with the total interrogative sentences which shows a declining F_0 pattern. Imperative sentences also show a declining F_0 contour and the range between some main peaks and valleys is larger and more abrupt than in declarative sentences.

The main features of global pitch changes can be summarized as:

Sentence modality	Global direction of pitch change	Mode of change	Degree of change
Declarative	declining	slurred	small
partial question	declining	slurred	medium
total question	rising	slurred	small
Imperative	declining	abrupt	large

Figures showing the pitch patterns corresponding to the sentence modalities under discussion are given in the appendix and the following sentences illustrate them:

Não antecipe o futuro. *(Do not anticipate the future.)*   [Fig. 1]
[nẽw̃ ẽtesipɪ u futuru]

Não antecipo o futuro. (*I don't anticipate the future.)*   [Fig. 2]
[nẽw̃ ẽtesipu u futuru]

O saldo é positivo? (*Is your bank balance favourable?)*   [Fig. 3]
[u sew sɑwdu ɛ pozitivu]

Quem antecipa o futuro? *(Who anticipates the future?)*   [Fig. 4]
[kẽj ẽtesipe u futuru]

REFERENCES

Albano, E. (1992) *Esboço de modelo neuropsicologicamente orientado para a aquisição de fonologia,* Anais do II Encontro Nacional sobre Aquisição de Linguagem, CEAAL/ PUC-RS, Porto Alegre.
_____ (1993) *Uma fonologia voltada para a implementação fonética,* I Encontro sobre Processamento de Língua Portuguesa Escrita e Falada, INESC/CLUL, Lisboa.
Aubergé, V. (1992) "Developing a structured lexicon for the synthesis of prosody", Bailly, G. Benoit, C. E Sawallis (eds.) *Talking Machines: Theories, Models and Designs,* Elsevier Science Publishers.
Bolinger, D. (1986) *Intonation and its Parts. Melody in Spoken English,* Edward Arnold Publishers Ltda.
Klatt, D. H. & Klatt, L. C. (1990) "Analysis, synthesis, and the perception of voice quality variations among female and male talkers", *Journal of the Acoustical Society of America,* 87, 820-857.
Larreur, D., Emerard, F., & Marty, (1990) *Linguistic and prosodic processing for a text-to-speech synthesis system,* European Conference on Speech Communication and Technology, Paris.
Madhukumar A.S., Rajendran, S. & Yegnanarayana, B. "Intonation component of a text-to-speech system for Hindi", *Speech and Language* 7, 283-301.
Pittam, J. (1987) "The long-term spectral measurement of voice quality as a social and personality marker: A review", *Language and Speech,* Vol.30, Part 1, Jan-Mar, 1987.
Quené, H. & René, K. (1992) "The derivation of prosody for text-to-speech from prosodic sentence structure", *Computer Speech and Language,* 6, 77-99.
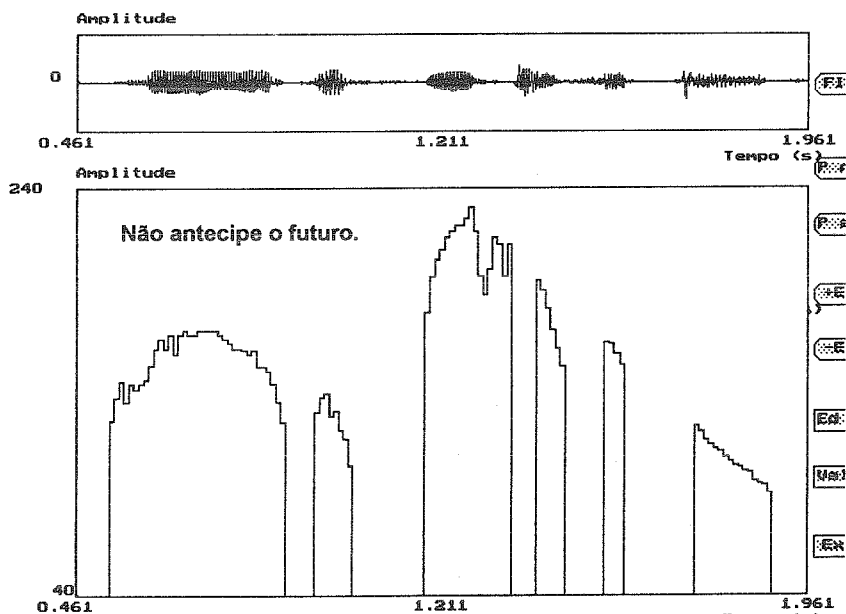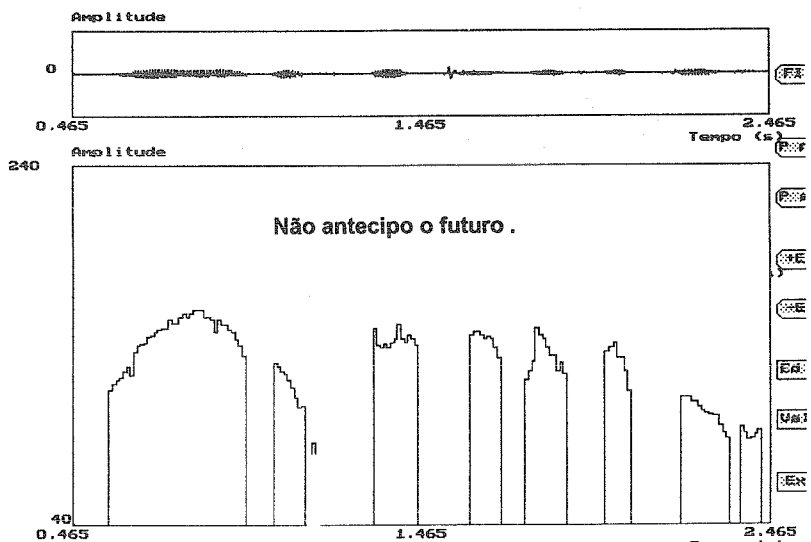
Appendix

Figure 1. Imperative sentence

Não antecipe o futuro.

Figure 2. Declarative sentence

Não antecipo o futuro .

O seu saldo é positivo ?

Figure 3. Total question

Quem antecipa o futuro ?

Figure 4. Partial question