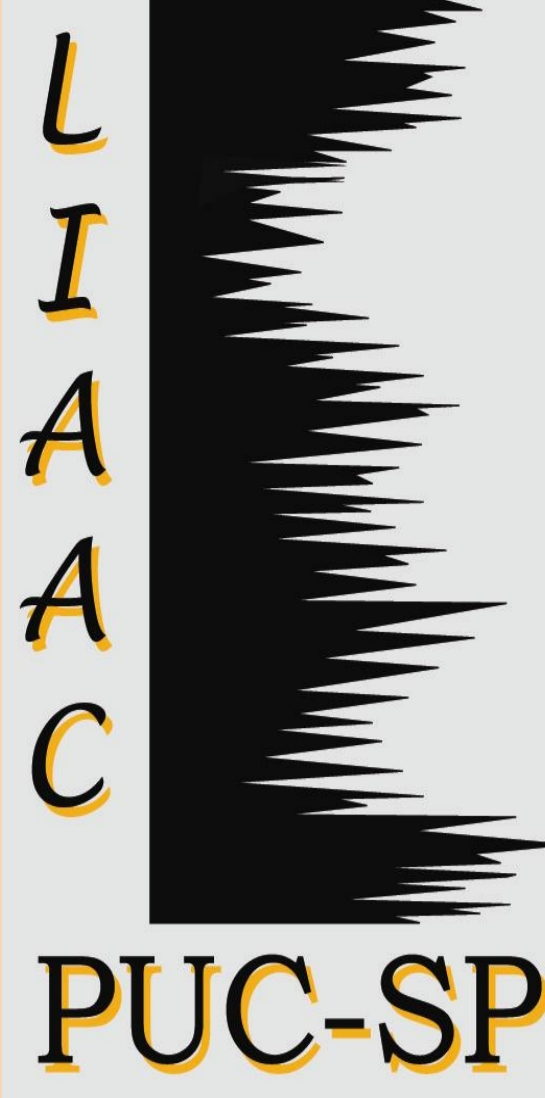# Reciting a sonnet: production strategies and perceptual effects

Sandra Madureira

Integrated Acoustic Analysis and Cognition Laboratory
Catholic University of São Paulo, São Paulo

madusali@pucsp.br

Abstract

This paper examines, from a dynamically-oriented approach, speaking strategies used by two professional speakers in reciting a sonnet. Spectrographic and perceptual analysis of prosodic and segmental elements are carried out. Uses of speaking strategies and their effects are contrasted to discuss relations between sound and sense.

Index Terms: sound and sense, speech expressivity, acoustic phonetic analysis, prosody, speaking styles.

Objective

The objective of this paper is to focus on the analysis of speaking strategy choices used by two speech professionals in reciting a sonnet in order to discuss how relations between sound and sense are built up.

Material

Commercially available recordings on CD of the "Soneto da Fidelidade" (Sonnet on Faithfulness) written by Vinicius de Moraes, a Brazilian diplomat, composer, song interpreter and poet recorded by an actor , identified in this work as S1 and an actress, identified as S2.

Methodological procedures

The following analysis procedures were carried out: perceptual evaluation, phonetic acoustic analysis and measurement of f0 values in syllabic rhymes and duration values of speech segments, syllables, V_V units (from vowel onset to vowel onset, consonant and silent pauses in this interval included).

Perceptual evaluation procedures

Voice qualities, pitch accents, use of allophones; rhythmical and speech rate variations and distribution of silent pauses were analyzed perceptually and based on inspection of acoustic data. Voice qualities were further classified by means of the VPAS.

Measurements procedures

Measurements of V-V units were taken into account due to their relevance as a psychoacoustic anchorage point as indicated in the phonetic literature on p-centers. Measurements of acoustic parameters were made by means of Praat. From the duration values obtained, the mean, standard deviation, z-score and smoothed z-score were calculated by the SG detector, a software developed by Plinio Barbosa from the University of Campinas whose script is transcribed in Barbosa (2006)

Pause classification criteria

To analyze the distribution of pauses, the notion of attractor was used. In this paper, the word "attractor" is used in the sense of a *locus* towards which some kind of behavior tends. The end of the stanza , the end of a verse or the boundaries between syntactic constituents in a verse, for instance, were considered to attract the placement of silent pauses, since speakers tended to produce silent pauses at these *loci*.

Four kinds of attractors were taken into account: pragmatic (PA); discursive (DA); syntactic (SA) and morph-syntactic (MSA). The ends of the stanza and of the verse, for instance, were considered discursive attractors and the boundaries between syntactic constituents a syntactic attractor.

Perceptual Evaluation Task procedures

Qualitative judgments on the performance of the two subjects were asked from a group of 30 adults. A protocol was built up for evaluation purposes. The judges listened to the recordings in a quiet room and filled in the protocol forms. The protocol had two items. The first was a question to investigate the overall impression caused on the listener. The second contained scalar degrees varying from 0 (absence) to 7 (presence in the highest degree) meant to be used to evaluate the conditions described by the following descriptors: sadness, anguish, softness and enthusiasm.

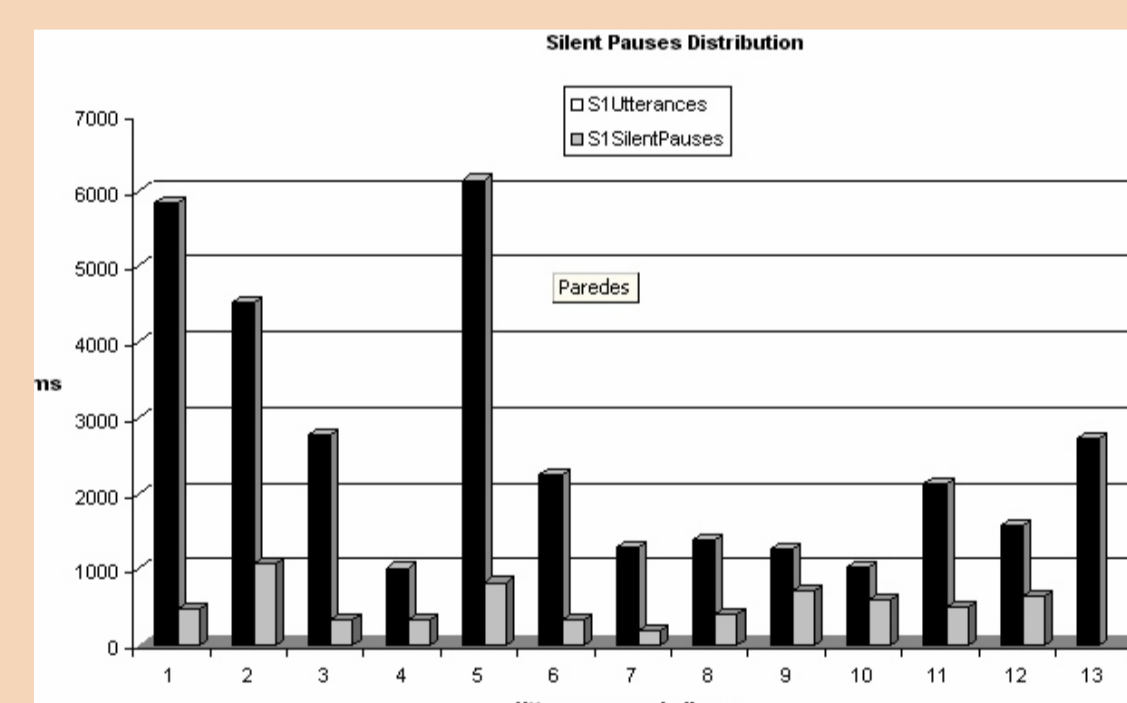## Results

1. The analysis of silent pauses

S1 produced less silent pauses (12) than S2 (24). S1 took 42 ms and S2 68 ms to recite the poem. S1 used mainly breathing pauses while S2 used pauses as a stylistic speaking strategy.

Duration in ms of S1's speech utterances in black and duration in ms of silent pauses in a light shade of gray.

Duration in ms of S2's speech utterances in black and duration in ms of silent pauses in a light shade of gray.
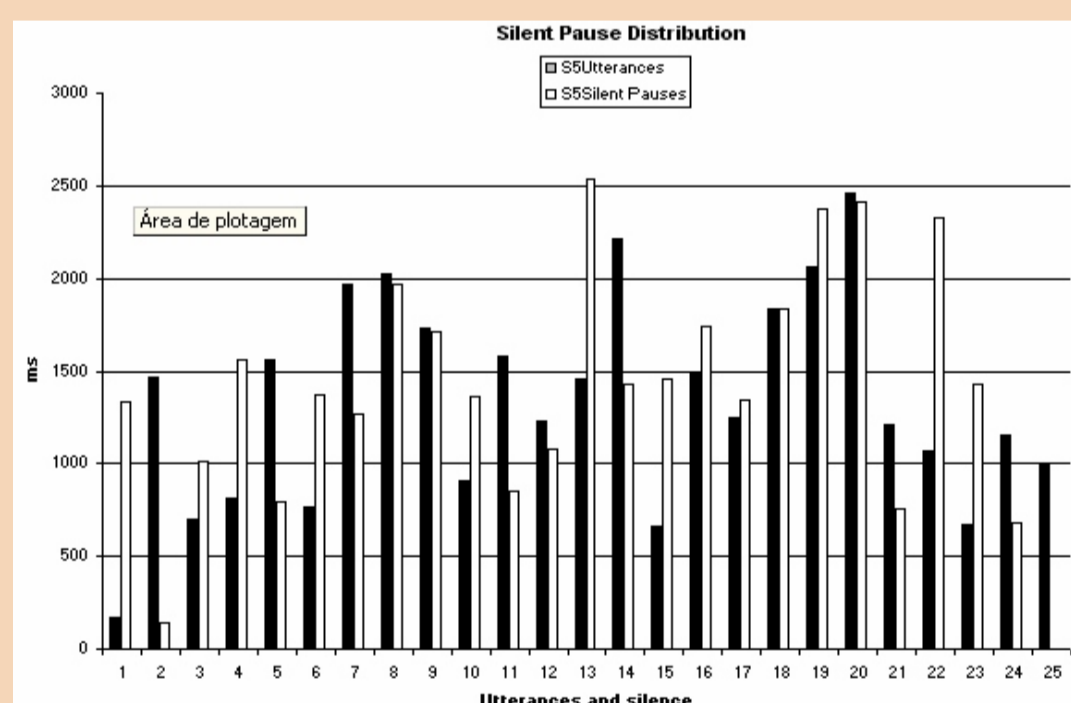
2. The analysis of voice qualities

S2 used *whispery voice* as a kind of permanent setting and *whispery creaky voice* as a kind of intermittent setting during the production of the sonnet and S1 used the setting of modal voice.

The circle refers to the vowel of the stressed syllable "cha" in the word "chama" (flame). The setting of creaky voice (VPAS) was found to characterize the S2's production.

3 . The analysis of f0 and duration contours
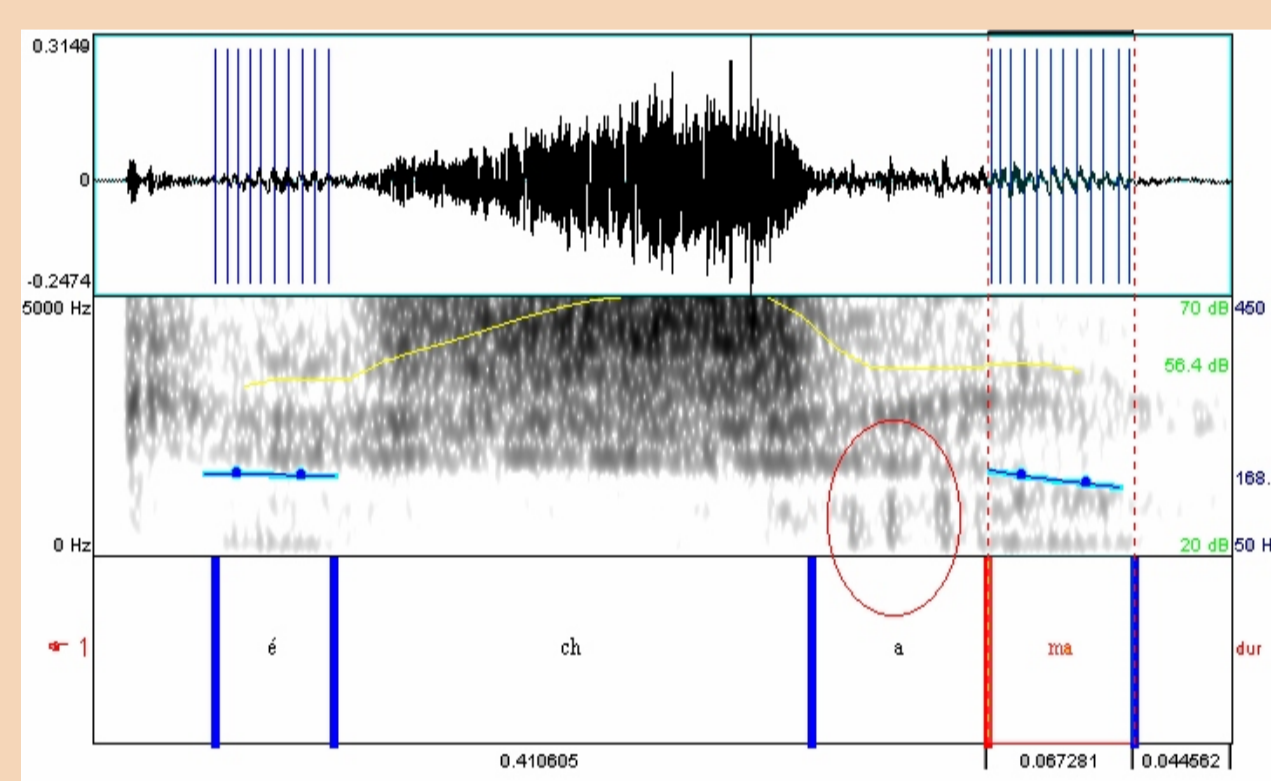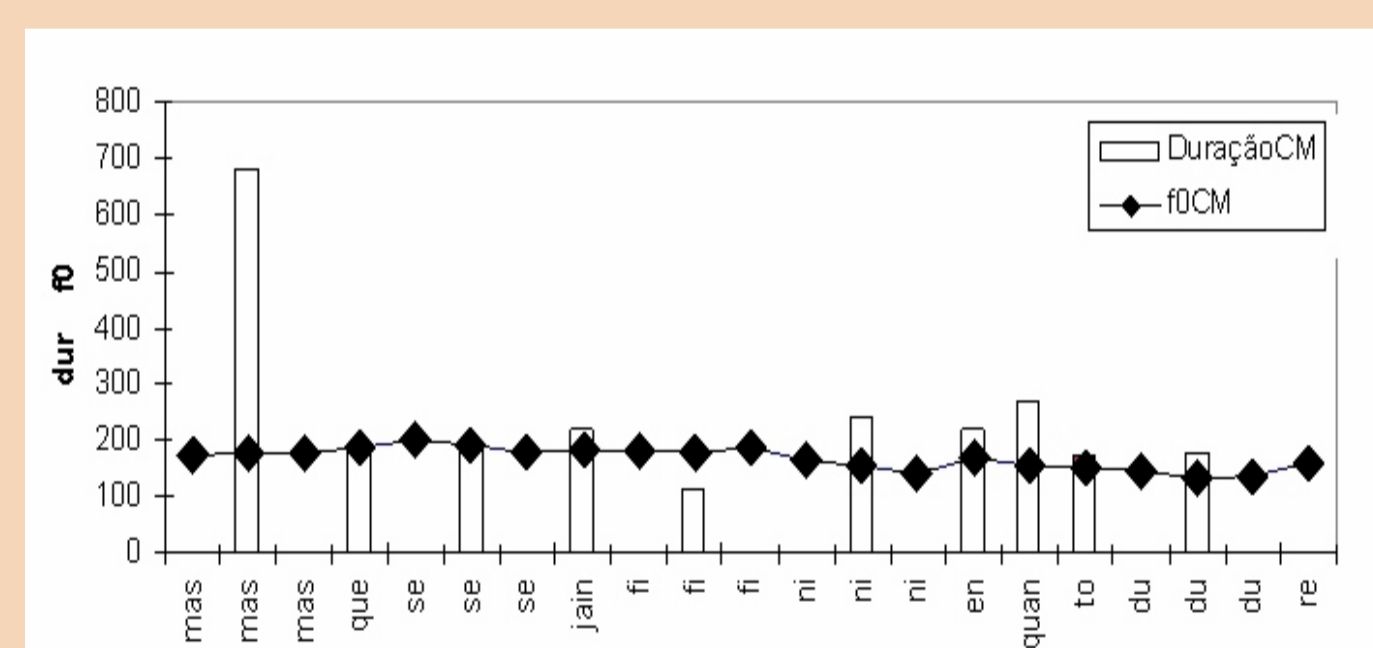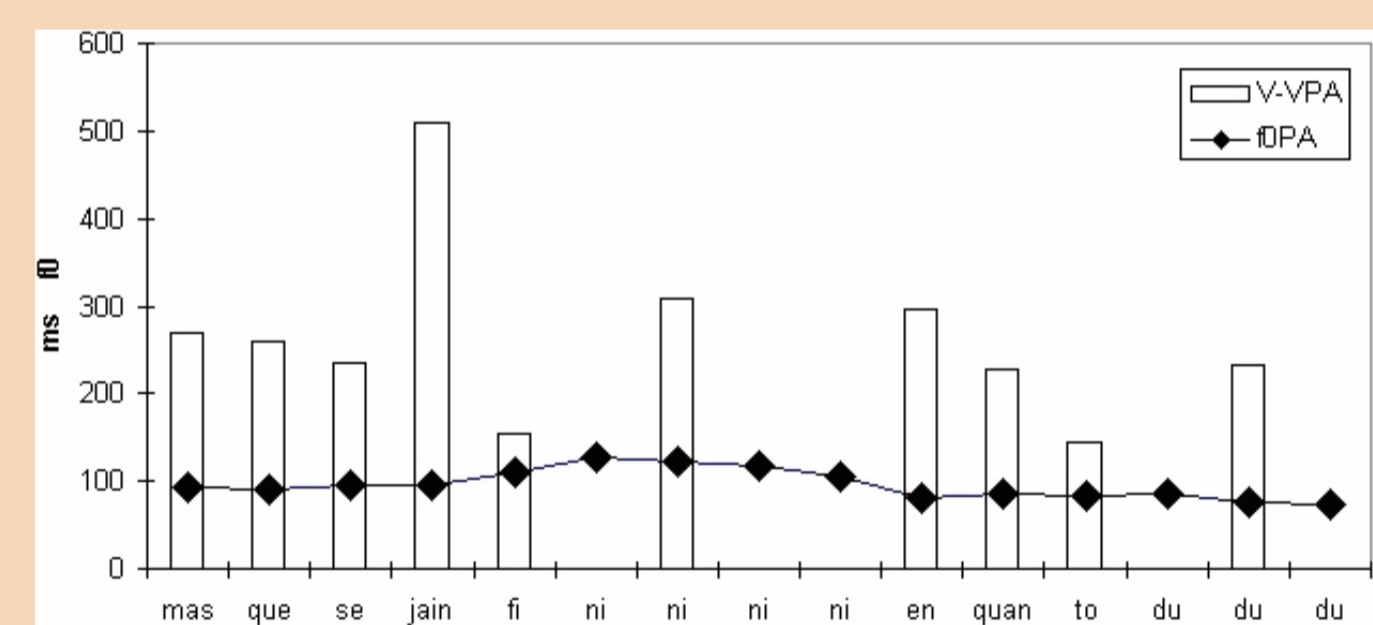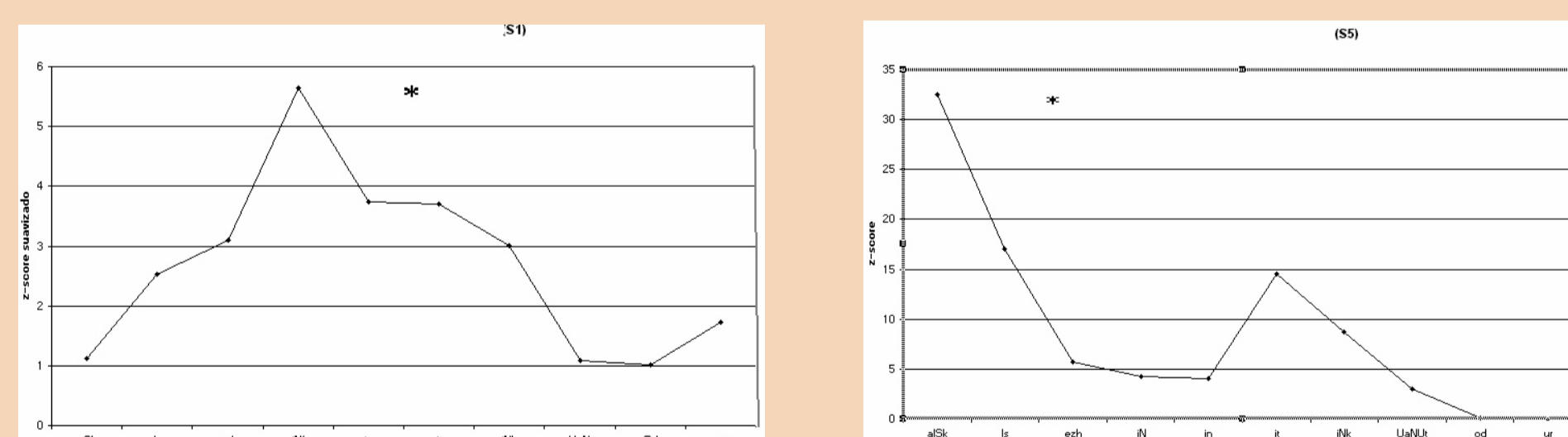
S1 varied intonation patterns as an expressive speaking strategy but S2 did not. Duration contours tended to differ.

F0 and duration contours of S1's production of the verse "mas que seja infinito enquanto dure" (upper graphic) and f0 and duration contour of S2's production of the same verse (lower graphic).

Duration maximum and f0 peak adjacency

Duration contour and f0 peak (*) in the production of the utterances "Mas (duration maxima in S2´s production) que seja infinito (duration maxima in S1´s production) enquanto dure"

4. Perceptual Evaluation Task

In answering the question about the overall impression caused by the speakers' oral interpretation, judges employed qualifiers such as "splendorous and gorgeous" when they listened to S1's performance and   "anxiety and grief" when they listened to S2's performance.
The results of the listeners' evaluation of affective states: - average values for S1: sadness (2.87); anguish (3.07); softness (3.13) and enthusiasm (4.67); -average values for S2: sadness (4.67); anguish (3.97); softness (4.57) and enthusiasm (2.27)Perceptual evaluation test

Discussion

S1 varied speaking strategies (salient pause distribution, pitch range, speech rate and kinds of speech patterns) a lot and often produced a climax effect while S2 used silent pauses, settings of *whispered voice and whispery creaky voice* and a narrow pitch range (usually from 140 Hz to 220 Hz) productively, which may have influenced the judges to report feelings of sadness and softness in a higher degree.

Conclusion

Comparing strategies used by professional actors in reciting poems, examining their effects on listeners and checking their text meaning production and performance intents is thought to be a promising undertaking since correlations between sound and sense can be made and productivity of strategies evaluated.
In this work, the speakers' prosodic choices (use of voice qualities and intonation patterns mainly) were found to be quite divergent and consequently impressed the listener in dissimilar ways.