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.

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 silenal pauses in this interval included).

Perceptual evaluation procedures
Voice qualities, pitch accent, use of allophones, rhythmic 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 VQAT.

Measurements procedures
Measurements of V-V' units were taken into account due to their relevance in psycholinguistic anchor points as indicated in the phonetic literature on p-continuities. 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 Pinto Barboza 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 stanzas, the end of a verse and 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 morphosyntactic (MSA). The ends of the stanzas and the boundaries between syntactic constituents constitute a syntactic attractor.

Results

1. The analysis of silent pauses
S1 produced less silent pauses (72) than S2 (96). 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.

2. The analysis of voice qualities
S2 used silenal pauses as a kind of permanent setting and silenal creasyor as a kind of a intermittent setting during the production of the sonnet and S1 used the setting of modal voice.

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.

Perceptual Evaluation Task

4. Perceptual Evaluation Task
In answering the question about the overall impression caused by the speakers’ oral interpretation, judges employed qualities such as “splendidious and gorgeous” when they listened to S1’s performance and “animity and grief” when they listened to S2’s performance. The results of the listeners evaluation of affective states—average values for S1: sadness (2.97); anger (0.75); softness (5.13) and enthusiasm (6.03); average values for S2: sadness (4.67); anger (1.97); softness (5.57) and enthusiasm (5.77)/Perceptual evaluation test.

Conclusion
The strategies used by professional actors in reciting poems, examining their effects on listeners and checking their next meaning production and performance intent 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 different and consequently impressed the listener in dissimilar ways.

Discussion
S1 varied speaking strategies (silent pause distribution, pitch range, speech rate and kinds of speech patterns) a lot and often produced a climate effect while S2 used silent pauses, settings of silenal voice and silenal creasyor 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.

SPEECH PROSODY 2008
Fourth Conference on Speech Prosody
May 6-9, 2008, Camperas, Brazil

Duration maximum and f0 peak adjacency
Duration duration and f0 peak adjacency
Duration maximum and f0 peak adjacency
Duration maximum and f0 peak adjacency
Duration maximum and f0 peak adjacency
Duration maximum and f0 peak adjacency