

Prosodic speech production and thematic segmentation

Antonis Botinis¹, Barbara Gawronska², Argyro Katsika¹ and
Dionisia Panagopoulou¹

¹*Dept of Linguistics, University of Athens*, ²*Dept of Languages, University of Skövde*

This study concerns tonal and temporal prosodic production in relation to thematic segmentation. A series of relatively short radio news broadcasts produced by female and male professional speakers with well-defined thematic alternations was analysed. The main results indicate: (1) thematic tonal onset and thematic initial stress are higher than thematic tonal offset and thematic final stress respectively (2) hypothematic tonal range is fairly embedded in hyperthematic tonal range and (3) hyperthematic silent pauses are significantly longer than hypothematic ones.

1. Introduction

The present study is within an experimental phonetics and computational linguistics research context with reference to the analysis and modeling of prosody (Botinis, Gawronska, Bannert & Sandblom, 2003). The long-term target of this research paradigm is the development and formalization of prosodic rules and prosodic modelling with reference to segmental, lexical, syntactic, semantic and discourse levels in speech production. Furthermore, the prosodic model is meant to have multiple applications, in the first place text-to-speech conversion but also machine translation as well as language education and medical rehabilitation applications. Several languages, such as German and Swedish, are investigated but here we will concentrate on the analysis of Greek.

Our analysis is based on a series of relatively short radio news broadcasts in Greek with reference to thematic segmentation as a function of tonal and temporal prosodic production. By prosodic tonal production we mean f_0 (voice fundamental frequency) realization at thematic initial and thematic final boundaries, i.e. thematic tonal onset and thematic tonal offset respectively, as well as f_0 realization at thematic initial stressed syllables and thematic final stressed syllables. By prosodic temporal production, on the other hand, we mean thematic speech durations as well as thematic silent pause durations distributed between major thematic units, referred to as hyperthemes and hypohemes (fairly corresponding topics and sub-topics, respectively, in much of the current literature).

Earlier research on this area (Botinis, 1995) has shown that there are structural tonal differences between thematic initialization and thematic finalization, which implies tonal correlates of thematic segmentation and thus thematic structuring. In the earlier analysis (i.e. Botinis, 1995) thematic silent pause distribution was not included in the data whereas the data itself was not subjected to statistical analysis. In this study, in addition to statistical analysis, we analyse considerably more data and we thus expect to have a fairly qualitative analysis and the very basic knowledge of prosodic strategies in thematic segmentation and boundary signaling of spoken discourse in general.

2. Experimental methodology

The speech material of this study consists of relatively short Greek radio news broadcasts, produced by professional media speakers, both male and female ones on separate occasions. The recorded speech material was written in standard Greek orthography and segmented in paragraphs, in accordance with traditional writing, i.e. the main themes of the news. Furthermore, the main themes, labeled as *hyperthemes*, were segmented into *hypotheses*, i.e. different aspects of the main theme. In the whole, the speech material consists of 10 news reports, produced by 5 female and 5 male speakers, with 58 hyperthemes and 127 hypotheses.

The speech analysis was carried out with the Wavesurfer software package at the University of Athens Phonetics Laboratory. In accordance with hyperthemes and hypotheses the following tonal measurements were taken: (1) thematic tonal onset, (2) thematic initial stress, (3) thematic final stress and (4) thematic tonal offset. Thematic tonal onset and thematic tonal offset were measured at the onset and offset of voiced segments respectively whereas thematic initial stress and thematic final stress were measured at the middle of the stressed syllable's vowel. Two temporal measurements were also taken i.e. (5) thematic speech production and (6) thematic silent pause.

Statistical analysis was carried out with the StatView software package with reference to thematic segmentation and prosodic measurements. T-tests, ANOVAs and correlation Z tests were carried out with reference to tonal and temporal measurements as a function of female vs. male speaker gender, hyper vs. hypo thematic category, and initial vs. medial vs. final thematic order.

3. Results

The main results of this study are presented in Figures 1-3. In Figure 1a, there is an anathetic tonal pattern from thematic tonal onset (female 225 Hz, male 141 Hz) to thematic initial stress (female 236 Hz, male 150 Hz) and a catathetic tonal pattern from thematic initial stress to thematic final stress (female 173 Hz, male 121 Hz) and thematic tonal offset (female 132 Hz, male 87 Hz). There are significant differences (t-test, 0.0001 level) as well as significant effects of speaker gender for all four measurement points (ANOVA, 0.0001 level). In Figure 1b, speaker gender has no noticeable effect on either thematic speech production durations or thematic silent pause durations.

In Figure 2a, the hypothematic tonal pattern is nested to the hyperthematic one but there was no significant effect of thematic category (at 0.05 level) except for the thematic tonal offset ($F=4.2$; $p<0.04$). Thus the hypothematic tonal onset (female 221 Hz, male 138 Hz) and initial stress (female 233 Hz, male 146 Hz) are lower than the respective hyperthematic tonal onset (female 235 Hz, male 148 Hz) and initial stress (female 244 Hz, male 161 Hz) whereas the reverse pattern is evident for hypothematic final stress (female 177 Hz, male 122 Hz) and tonal offset (female 136 Hz, male 89 Hz) which are higher than the respective hyperthematic final stress (female 167 Hz, male 122 Hz) and tonal offset (female 124 Hz, male 86 Hz). In Figure 2b, hyperthematic speech production durations are longer for female speakers than male ones (female 18.104 ms, male 14.730 ms) whereas hypothematic speech production durations have the opposite tendency (female 5.275 ms, male 5.785 ms). Hyperthematic silent pause durations between female and male speakers are fairly even (female 1.089 ms, male 1.122 ms) whereas hypothematic speech production durations are considerably longer for male speakers than female ones (female 510 ms, male 730 ms). Thematic category had a

significant effect on both speech production ($F=59.4$; $p<0.0001$) and silent pause durations ($F=8.3$; $p<0.004$) whereas the effect of speaker gender was not significant (at 0.05 level).

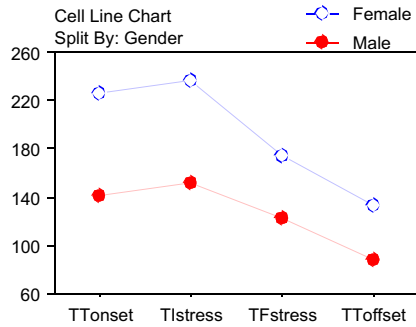


Figure 1a. Thematic tonal onset (TTonset), initial stress (Tlstress), final stress (TFstress) and tonal offset (TToffset) as a function of female vs. male speaker gender.

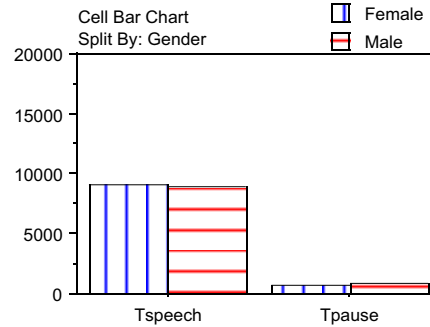


Figure 1b. Thematic speech production durations (Tspeech) and silent pause durations (Tpause) as a function of female vs. male speaker gender.

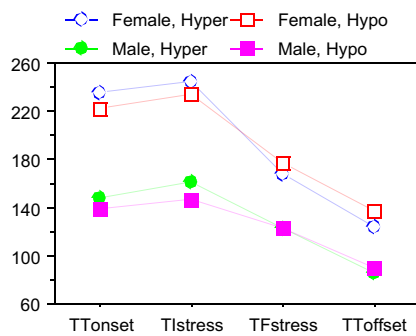


Figure 2a. Thematic tonal onset (TTonset), initial stress (Tlstress), final stress (TFstress) and tonal offset (TToffset) as a function of female vs. male speaker gender and hyper vs. hypo thematic category.

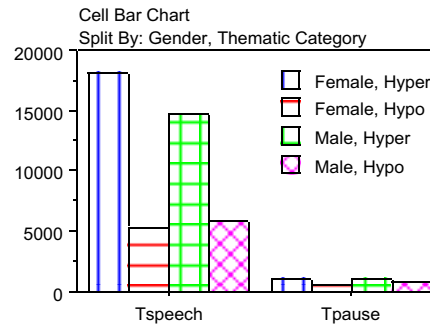


Figure 2b. Thematic speech production durations (Tspeech) and silent pause durations (Tpause) as a function of female vs. male speaker gender and hyper vs. hypo thematic category.

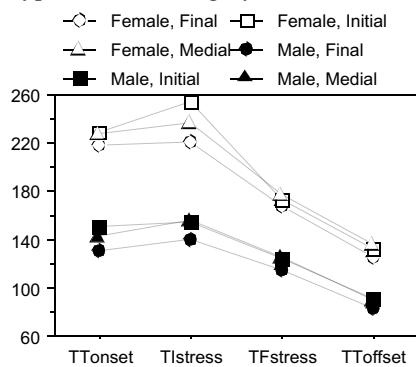


Figure 3a. Thematic tonal onset (TTonset), initial stress (Tlstress), final stress (TFstress) and tonal offset (TToffset) as a function of female vs. male speaker gender and initial vs. medial vs. final thematic order.

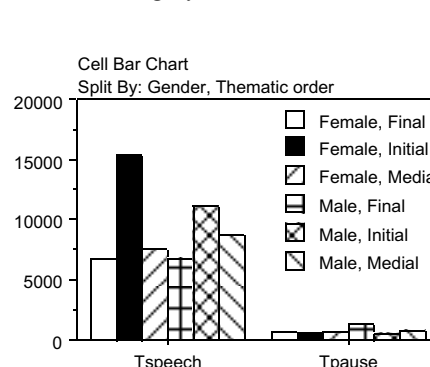


Figure 3b. Thematic speech production durations (Tspeech) and silent pause durations (Tpause) as a function of female vs. male speaker gender and initial vs. medial vs. final thematic order.

In Figure 3a, the tonal pattern of the thematic final order is generally lower than that of initial and medial order. However, thematic order reached significance only for thematic initial stress ($F=4.0$; $p<0.01$) but not for thematic tonal onset, thematic final stress and thematic tonal offset (at 0.05 level). Scheffe's post hoc test showed significant effects for final and initial order ($p<0.02$) as well as for final and medial order ($p<0.001$) but not for initial and medial order. In Figure 3b, the thematic speech production durations of initial order are significantly longer than that of medial and final order ($F=3.7$; $p<0.02$) but Scheffe's post hoc test showed significant effects only for female speaker gender. Thematic silent pause durations were marginally significant ($F=2.9$; $p<0.05$) and this effect of thematic order was confined to male speaker gender ($F=3.8$; $p<0.02$). Correlation Z test did not show any significance with reference to thematic speech production and thematic silent durations.

4. Discussion and conclusions

Although most of the results deserve a fair share of attention we will concentrate on main issues. First, although the intrinsic gender tonal productions are generally kept, they are minimized at thematic tonal offset. In another study (Bannert et al., 2003) there was no significant gender effect at thematic tonal offset, which implies that female speakers may reach the edge of their tonal range for thematic segmentation and discourse structuring. Thematic speech production and thematic silent durations, on the other hand, do not show any divergence between female and male speakers. Second, hyperthematic and hypothematic categories are distinct with regards to thematic tonal offset, i.e. final tonal boundary, as well as silent pause durations. The latter point should be emphasized, as the distribution of silent pauses seems rather categorical, i.e. either hyperthematic or hypothematic, which is associated with different types of pauses and hence pause durations rather than degree. In the present material initial thematic categories were significantly longer than medial and final ones but this had hardly any effect on silent pause durations. There seems thus that thematic tonal offset in combination with tonal resetting, as evidenced by thematic tonal onset of succeeding thematic unit, as well as silent pause duration are primary prosodic cues for thematic segmentation as well as thematic categorisation. Further prosodic cues for thematic segmentation and categorization (see Botinis, 1992), such as initial thematic accent or tonal onset variability, may be found in other types of news materials or spontaneous speech.

5. References

- Bannert, R., Gawronska, B. Katsika, A. & Sandblom, E. (2003) Discourse structure and prosodic correlates. In *Proceedings of the fifteenth international congress of phonetic sciences*, Barcelona, Spain (to appear).
- Botinis, A. (1992) Accentual distribution in Greek discourse. *Travaux de l'Institut de Phonétique d'Aix*, vol. 14, 13-52.
- Botinis, A. (1995) Tonal correlates of discourse structure. In *Proceedings of the thirteenth international congress of phonetic sciences*, Stockholm, Sweden, pp. 366-369.
- Botinis, A., Gawronska, B., Bannert, R. & Sandblom, E. (2003) Crosslinguistic analysis and modeling of prosody. In *Proceedings of the fifteenth international congress of phonetic sciences*, Barcelona, Spain (to appear).