

Discourse Structure and Prosodic Correlates

Robert Bannert[†], Antonis Botinis[‡], Barbara Gawronska^{††} and Argyro Katsika[‡]

[†] University of Umeå, Sweden

[‡] University of Athens, Greece

^{††} University of Skövde, Sweden

E-mail: Robert.Bannert@ling.umu.se, abotinis@phil.uoa.gr, barbara.gawronska@isp.his.se

ABSTRACT

This investigation is part of a larger study with reference to the analysis of discourse structure from a crosslinguistic perspective in a computational context. TV news in Greek will be presented in this paper with main emphasis on segmentation of thematic units and prosodic correlates. Both local and global tonal cues as well as silent pauses are evident in segmentation of thematic units with reference to discourse structure.

1. INTRODUCTION

The present investigation is part of a larger study where a crosslinguistic acoustic analysis of discourse structure and prosodic correlates is carried out [1]. Several languages are studied, including English, German, Greek and Swedish but, in this paper, Greek data will be presented as an exemplification of our research paradigm. Discourse structure mainly refers to forms, relations and functions of thematic units whereas prosodic correlates refer to tonal, temporal and dynamic prosodic cues. In this paper we concentrate on thematic segmentation in relation to tonal and temporal prosodic correlates.

Tonal cues are associated with a wide-range of speech units such as lexical, syntactic and discourse ones. The tonal cues may be at a local domain as well as at a fairly global domain with well-defined tonal forms along with tonal alignment and tonal range variations. Silent pauses may also be associated with variable speech units, especially syntactic and discourse ones. Prosodic structures at e.g. lexical level may have both different forms and functions across different languages (e.g. [2], [3], [4], [5]). What is the case at discourse level? How do languages with different prosodic structures use prosodic cues in discourse marking such as segmentation?

There is considerable knowledge in the prosodic structure of different languages, including Greek, which is rapidly accelerating with reference to prosodic distinctions and cues in relation to main linguistic components (see [2], [5], [6]). The primary aim of this investigation is thus to investigate discourse prosodic cues which, in addition to prosodic research, are related to text linguistics and linguistic theory in general.

2. ANALYSIS PROCEDURES

The speech material in the framework of this study consists of TV news as well as radio and spontaneous dialogues in several languages. In the present analysis, however, we concentrate on the analysis of Greek TV news produced by a male and a female speaker on two separate occasions. Both speakers are relatively young, i.e. in their thirties, with a typical media accent.

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, by the third author. In the whole, the speech material consists of 14 hyperthemes and each hypertheme consists of 3-4 hypotheses.

The speech analysis was carried out with the waveserfer software package at the Phonetics Laboratory, University of Athens, by the third author. With reference to thematic structure the following tonal (F0) measurements were taken according to hyperthemes and hypotheses: (1) thematic tonal onset, (2) thematic first stressed syllable, (3) thematic last stressed syllable and (4) thematic tonal offset. Two measurements were also taken with reference to temporal patterns, i.e. (5) thematic speech duration and (6) thematic pause duration.

Statistical analysis was carried out with the StatView statistical package with reference to thematic segmentation and prosodic measurements. Comparison t-tests, ANOVAs, as well as correlation z-scores were carried out with reference to tonal and temporal measurements as a function of thematic categories, i.e. hyperthemes vs. hypotheses, thematic order, i.e. initial vs. medial vs. final, as well as gender, i.e. female vs. male.

In the framework of the present analysis, spoken discourse is assumed to consist of thematic units, which are basic speech communication units with a fairly well-organised prosodic structure and segmentation correlates. In addition to other linguistic cues, such as lexical and syntactic ones, prosody is assumed to have a distinctive function in thematic segmentation, which is a basic aspect of discourse structuring.

3. THEMATIC SEGMENTATION

Figures 1-5 show thematic segmentation and prosodic correlates in accordance with the present methodology tonal and temporal measurements (see section 2), i.e. *thematic tonal onset, thematic first stressed syllable, thematic last stressed syllable, thematic tonal offset, thematic speech duration and thematic pause duration.*

Figure 1 shows the tonal measurements of the total speech material (see above), i.e. both female and male speakers, in accordance with thematic segmentation. There is an anathetic pattern from thematic onset (169 Hz) onto thematic first stressed syllable (212 Hz) whereas the opposite pattern, i.e. a catathetic pattern, from thematic first stressed syllable (212 Hz) to thematic last stressed syllable (138 Hz) and even deeper to thematic offset 106 Hz). Paired t-tests showed highly significant differences among all four points of measurements ($p < 0.001$). A strong tonal lowering is thus evident from left to right with reference to both tonal tops and tonal bottoms of the analysed material.

Figure 2 shows the tonal measurements of the speech material (see above) as a function of gender, i.e. female vs. male. The effect of gender was highly significant for the first three measurement points ($p < 0.0001$), i.e. thematic tonal onset (female 255 Hz vs. male 129 Hz), thematic first stressed syllable (female 292 Hz vs. male 174 Hz) and thematic last stressed syllable (178 Hz vs. male 119 Hz). The last measurement, however, i.e. thematic tonal offset, did not reach a significant level (at 0.05), despite a noticeable difference for approximately 20 Hz (female 121 Hz vs. male 99 Hz). The gender distinction is thus evident with reference to tonal production, but for major linguistic segmentation purposes.

Figure 3 shows the tonal measurements of the speech material (see above) as a function of thematic category, i.e. hyperthematic vs. hypothematic. The effect of thematic category did not reach a significant level (at 0.05), although there are noticeable differences between these two categories. The thematic tonal onset (hyperthematic 185 Hz vs. hypothematic 165 Hz) and the thematic first stressed syllable (hyperthematic 221 Hz vs. hypothematic 209 Hz) of the hyperthemes are slightly higher than that of the hypothemes whereas the opposite structure is taking place for the thematic last stressed syllable (hyperthematic 118 Hz vs. hypothematic 143 Hz) and the thematic tonal offset (hyperthematic 83 Hz vs. hypothematic 112 Hz). In general, the hypothematic category seems to have a tendency for a nested to hyperthematic category tonal structure but, most probably, the amount of the present data does not suffice for a significant effect.

Figure 4 shows the tonal measurements of the speech material (see above) as a function of thematic order, i.e. initial vs. medial vs. final. The thematic tonal onset (initial 183 Hz vs. medial 165 Hz vs. final 166 Hz) and the thematic first stressed syllable (initial 218 Hz vs. medial 211 Hz vs. final 208 Hz) of initial themes have a tendency to be slightly higher than that of medial and final ones

whereas the thematic last stressed syllable (initial 138 Hz vs. medial 145 Hz vs. final 120 Hz) and the thematic tonal offset (initial 105 Hz vs. medial 115 Hz vs. final 87 Hz) have a tendency to be slightly lower than that of medial and initial ones. However, despite these noticeable differences, the effect of thematic order does not reach a significant level (at 0.05). There is thus a tendency for a left-to-right tonal lowering even for thematic order but most probably much more data is required for verification of this effect.

Figure 5 shows the duration of thematic speech as well as thematic pause duration as a function of thematic category, i.e. hyperthemes and hypothemes. The effect of thematic category is highly significant ($p < 0.0001$), i.e. the thematic pause duration of the hyperthemes (1607 ms) is considerably longer than that of hyperthemes (409 ms). It is thus most evident that a division of thematic structure into two major categories, i.e. hyperthemes and hypothemes has a temporal realisation as a strong prosodic and thus thematic boundary.

Figure 6 shows thematic tonal offset and thematic pause duration as a function of hyperthematic and hypothematic category. There is a negative correlation between thematic tonal offset and thematic pause duration, according to which lower tonal values are correlated with longer pause duration values and, vice versa, higher tonal values are correlated to shorter pause duration values. The former correspond to hyperthematic category whereas the latter to hypothematic one. This correlation did reach significance ($p < 0.0001$, see bellow). Thematic tonal offset and thematic pause duration thus seem to have a compensatory distribution for thematic segmentation.

The tonal and temporal measurements of the speech material is subjected to correlation analysis (Z test), the main results of which are as follows:

1. Significant correlation between thematic tonal onset and thematic speech duration ($Z=3.3$, $p < 0.0008$).
2. Significant correlation between thematic tonal onset and thematic pause duration ($Z=4.0$, $p < 0.0001$).
3. Significant correlation between thematic first stressed syllable and thematic speech duration ($Z=3.0$, $p < 0.002$).
4. Significant correlation between thematic first stressed syllable and thematic pause duration ($Z=4.0$, $p < 0.0001$).
5. Non-significant correlation between thematic tonal offset and thematic speech duration (at 0.05 level).
6. Significant correlation between thematic tonal offset and thematic pause duration ($Z=4.9$, $p < 0.0001$).

In summary, there are significant correlations between key measurements of the analysed material, which indicate a tonal and temporal interplay for thematic segmentation.

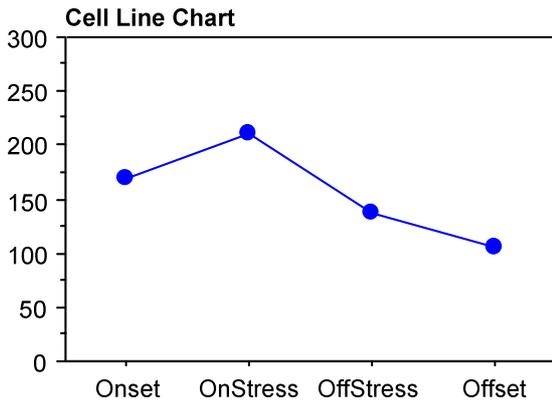


Figure 1. Tonal measurements (in Hz) of thematic onset (Onset), thematic first stressed syllable (OnStress), thematic last stressed syllable (OffStress) and thematic offset (Offset).

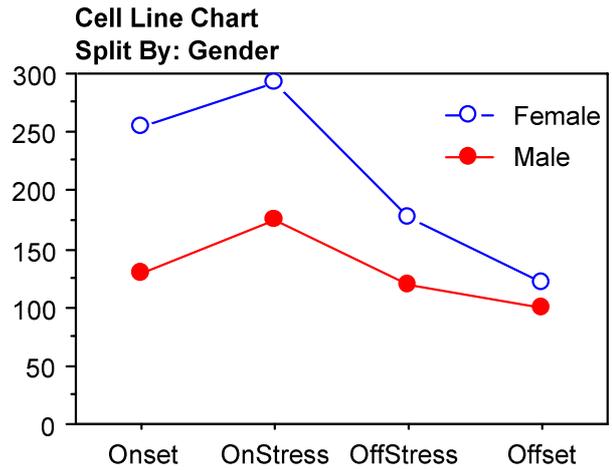


Figure 2. Tonal measurements (in Hz) of thematic onset (Onset), thematic first stressed syllable (OnStress), thematic last stressed syllable (OffStress) and thematic offset (Offset) as a function of gender (female vs. male).

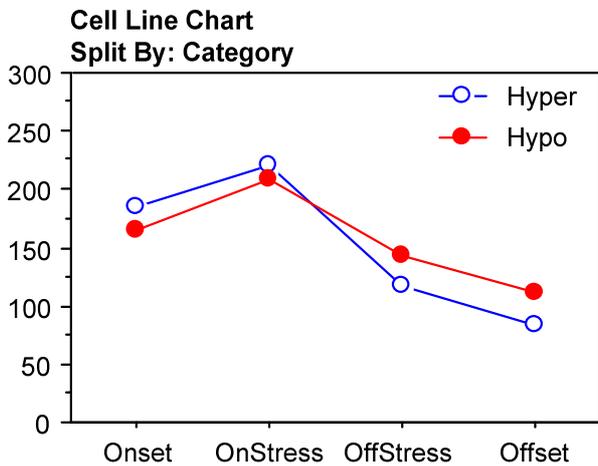


Figure 3. Tonal measurements (in Hz) of thematic onset (Onset), thematic first stressed syllable (OnStress), thematic last stressed syllable (OffStress) and thematic offset (Offset) as a function of thematic category (hypertheme vs. hypotheme).

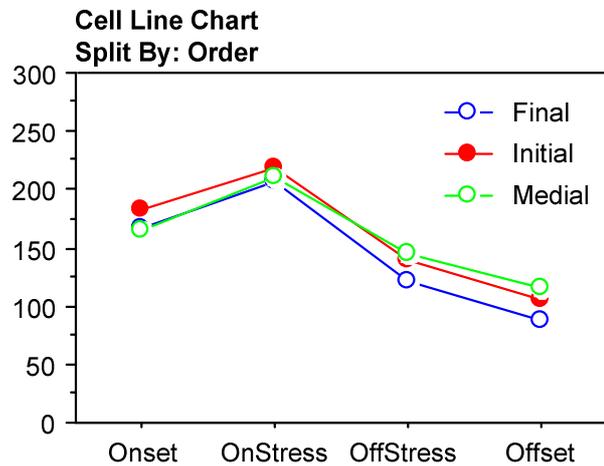


Figure 4. Tonal measurements (in Hz) of thematic onset (Onset), thematic first stressed syllable (OnStress), thematic last stressed syllable (OffStress) and thematic offset (Offset) as a function of thematic order (initial vs. medial vs. final).

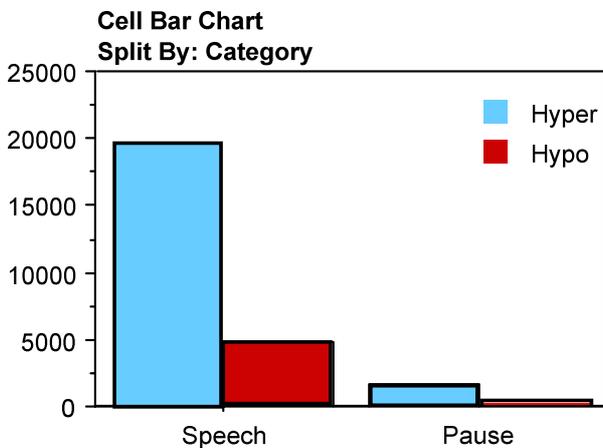


Figure 5. Thematic speech (in ms) and silent pause durations as a function of thematic category (hypertheme vs. hypotheme).

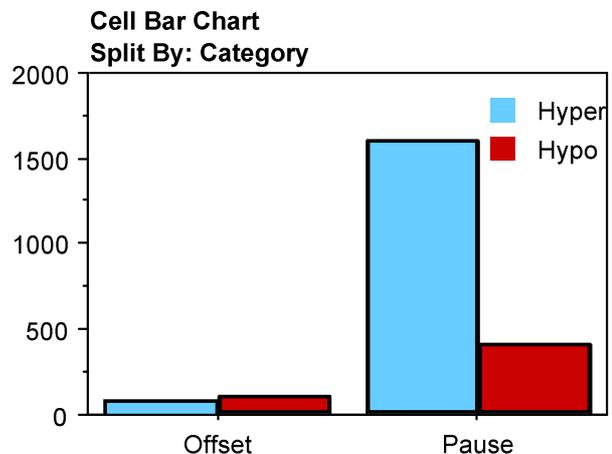


Figure 6. Thematic tonal offset (in Hz) and silent pause durations (in ms) as a function of thematic category (hypertheme vs. hypotheme).

4. DISCUSSION AND CONCLUSIONS

The main results of the present analysis are based on TV news and are thus biased towards this type of materials (cf. [7], [8], [9], [10]).

First, thematic tonal onset and thematic first stressed syllable are higher than the thematic tonal offset and the thematic last stressed syllable respectively, thus resulting in a global rising-falling tonal pattern.

Second, both female and male speakers reach the bottom of their respective tonal-range at thematic offset where gender differences are minimised, although no tonal divergence as a function of gender is evident at other measurement points.

Third, thematic categories, i.e. hyperthemes and hypothemes, as well as thematic order, i.e. initial, medial and final, do not show significant differences in this analysis although there is a tendency for larger tonal range variations of hyperthemes, with reference to both thematic onset and thematic offset.

Fourth, thematic pause durations show a distinct pattern, according to which hyperthemes do have considerably longer pause durations in comparison to that of the hypothemes.

The tonal structure of the present analysis is in general agreement and corroborates earlier results on thematic segmentation and tonal correlates of Greek news reports [9]. A significant finding of the present analysis, however, concerns the tonal minimalisation between female and male speakers at thematic tonal offset (i.e. final boundaries). This implies that tonal lowering, which may even reach the speaker's bottom tonal range, is a strong prosodic correlate of thematic finality. To this end, there seems to be a meaningful effort on behalf of a female speaker for thematic finality and thematic segmentation. On the other hand, there is a standard tonal difference at all other measurement points, according to which female tonal realisation is much higher than male one.

In general, speakers tend to reach the edge of their physiological constraints for linguistic purposes. In addition to thematic finality tonal realization outlined above, the phenomenon of "pitch-concord", i.e. an inter-speaker adjustment at approximately the same level, even when speakers of different age/gender are involved, has been reported in Greek [11]. Normally, physiologically determined tonal differences between the speakers should be expected.

The overall tonal pattern of the present speech material shows fairly good similarities to the tonal pattern of laboratory speech materials, i.e. an overall tonal declination from right to left. In addition to relatively short utterances, the declination effect has also been reported for longer texts (see [12]).

REFERENCES

- [1] A. Botinis, B. Gawronska, R. Bannert and E. Sandblom, "Crosslinguistic analysis and modelling of prosody", *Proc. 15th ICPHS*, Barcelona, Spain (to appear).
- [2] A. Botinis, *Stress and Prosodic Structure in Greek*, Lund University Press, 1989.
- [3] A. Botinis, R. Bannert and M. Tatham, "Contrastive tonal analysis of focus perception in Greek and Swedish", in A. Botinis (ed.), *Intonation: Analysis, Modelling and Technology*, pp. 97-116, Dordrecht: Kluwer Academia Publishers, 2000.
- [4] A. Botinis, B. Granström and B. Möbius, "Developments and paradigms in intonation research", *Speech Communication*, vol. 33, 263-296, 2001.
- [5] A. Di Cristo and D. Hirst, 1998. *Intonation Systems*, Cambridge University press.
- [6] M. Rossi, "Intonation: past, present, future", in A. Botinis (ed.), *Intonation: Analysis, Modelling and Technology*, pp. 13-52, Dordrecht: Kluwer Academia Publishers, 2000.
- [7] B. Gawronska and D. House, "Information Extraction and text generation of news reports for a Swedish-English bilingual spoken dialogue system", *ICSLP '98*, Sydney, Australia, vol. 4, pp. 1139-1142, 1998.
- [8] B. Gawronska, D.I. Jakobsson and J. Anttila, "Mental spaces, space builders and bilingual summarization of news reports", *8th International Conference on Theoretical and Methodological Issues in Machine Translation (TMI 99)*, Chester, England, pp. 1-10, 1999.
- [9] A. Botinis, "Tonal correlates of discourse structure", *Proc. 13th ICPHS*, Stockholm, Sweden, vol. 2, pp. 366-369, 1995.
- [10] R. Bannert, "Variations in the perceptual modelling of macro-prosodic organization of spoken Swedish: prominence and chunking". Reports from the Department of Phonetics, Umeå University, *PHONUM* 3, pp. 31-53, 1995.
- [11] A. Botinis, "Accentual distribution in Greek discourse". *Travaux de l'Institut de Phonetique d'Aix*, vol. 14, pp. 13-52, 1992.
- [12] N. Grønnum, "Intonation in Danish", in A. Di Cristo and D. Hirst (eds.), *Intonation Systems*, Cambridge University press, pp. 131-151, 1998.