

Produced pauses, perceived pauses and thematic units

Antonis Botinis^{1,2}, Aikaterini Bakakou-Orphanou¹, Harris Tafroglou¹ and Anastasia Christopoulou¹

¹Department of Linguistics, University of Athens

²Department of Communication and Information, University of Skövde

Abstract

In this study we investigate produced and perceived pauses in relation to thematic units of Greek radio news, defined by both listeners and readers. The results indicate: (1) perceived pauses are more than both produced pauses and thematic units; (2) the number of thematic units by listeners and readers does not differ substantially; (3) an agreement of $\geq 80\%$ among listeners and readers results in a substantial drop of thematic units; (a) neither short and long pause nor hypertheme and hypoheme categories are adequately perceived.

Introduction

This study is within the framework of a multi-factor research paradigm in text and dialogue linguistics. We examine lexical syntactic, semantic and prosodic factors in an integrated experimental framework of experimental phonetics, theoretical linguistics and computational linguistics. Examples of our research is reported in Botinis, Gawronska, Katsika and Panagopoulou (2003), Botinis, Gawronska, Bannert and Sandblom (2003), Botinis, Bakakou-Orphanou, Tsimokas and Patsalou (2004).

In this paper, we will report three experiments with reference to short radio news in Greek. First, an acoustic analysis of produced pauses. Second, an auditory analysis of perceived pauses. Third, both a listening and reading analysis of thematic segmentation. In accordance with the experiments three main questions are addressed: (1) what is the relation between produced pauses and perceived pauses? (2) what is the relation between the durations of produced pauses and the perceived length of the respective pauses? (3) what is the relation between produced/perceived pauses and thematic units segmentation?

Our research intends to the accumulation of basic linguistic knowledge in text and dialogue areas, which is expected to lead to the development of powerful models with potential applications in language technology areas such as text and speech parsing, information extraction, speech synthesis and automatic translation.

Experimental methodology

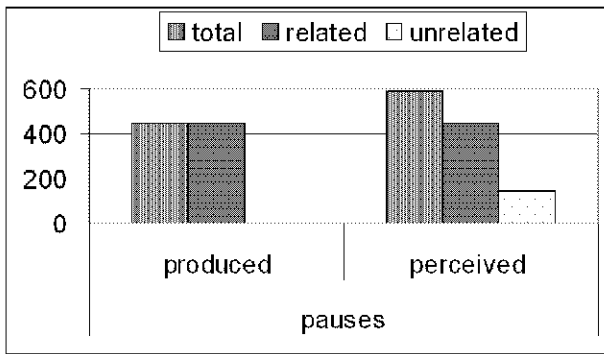
The speech material under experimentation consists of twenty (20) one-minute radio news in Greek, half of which were produced by female and half by male speakers. The speakers are media professionals with standard Greek pronunciation and the speech material was directly recorded and stored on a PC computer disc for further processing with reference to three experiments carried out in this study.

In the first experiment, all produced pauses, both silent and breath ones, were measured in milliseconds with the Waveserfer software package. In the second experiment, one group of ten Greek students of the Linguistics Department at the University of Athens listened to the speech material as much as they liked and classified the perceived pauses as either short or long ones, on their own individual subjective basis. In the third experiment, another group of ten Greek students of the Linguistic Department at the University of Athens segmented the speech material into hyperthemes (i.e. topics) and hypohemes (i.e. subtopics).

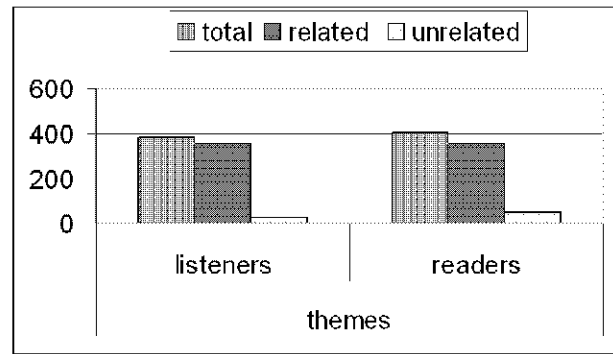
The third experiment involved two tasks, a listening and a reading one. In the listening task, the speech material was written in continuous text in standard Greek orthography with all punctuation marks but paragraph ones. Two experimental groups of ten postgraduate Greek students at Athens University Linguistics Department each, who had at their disposition either both text and speech or only speech, were asked to segment the speech material in hyperthemes and hypohemes, in accordance with both text and speech or only text linguistic information.

Results

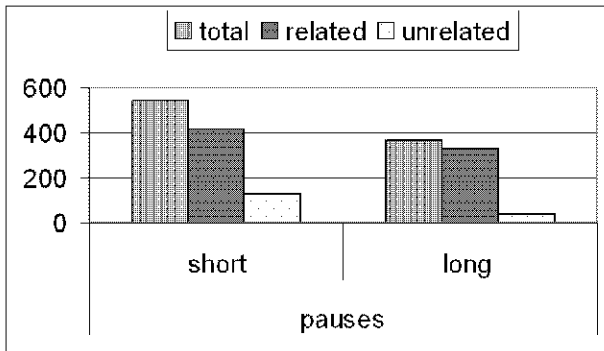
The results of the experiments are presented in Figures 1-3. It should be noted that the total number of pauses and themes is based on at least one listener's single perceptual response. Agreements among listeners from 10% to 100% (indicated 1-10) is also calculated and presented in Figures.



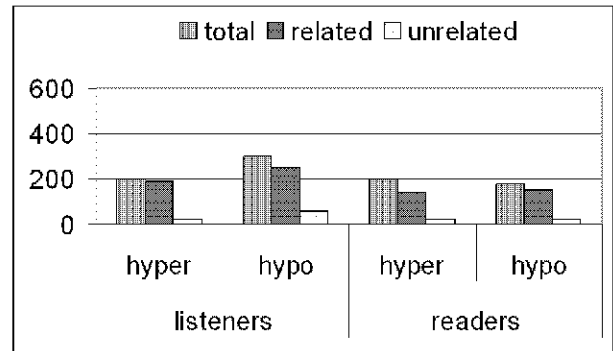
(a)



(c)



(b)



(d)

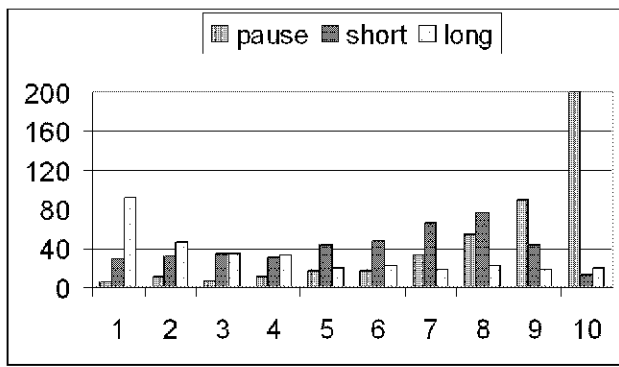
Figure 1. Frequencies of the production and perception experiments: (a) produced and perceived pauses, (b) perceived short and long pauses, (c) thematic segmentation by listeners and readers, (d) hyperthematic (hyper) and hypothematic (hypo) segmentation by listeners and readers. Total frequencies well as related and unrelated to produced pauses frequency are indicated.

Figure 1a shows the total number of produced and perceived pauses as well as the related and unrelated to each other pauses is presented. The produced pauses are more than the perceived pauses. In total, the produced pauses are 444 and the perceived pauses are 587. Furthermore, all 444 produced pauses are perceived (by at least one listener) and thus 444 perceived pauses are related to produced pauses whereas the remaining 143 perceived pauses are auditory pauses with no relation to produced physical pauses.

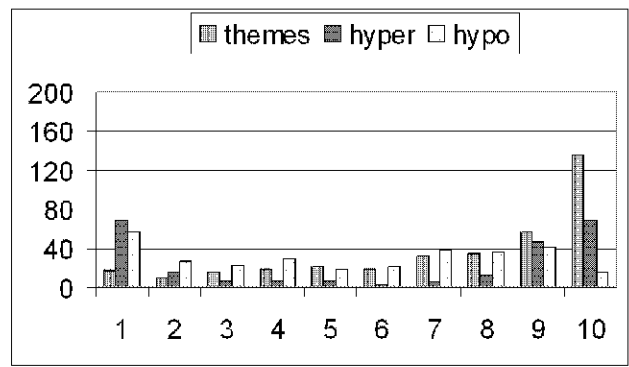
Figure 1b shows the total number of perceived short and long pauses as well as related and unrelated perceived pauses to produced ones. The perceived short pauses in total are more than the long ones. The long pauses, on the other hand, are more related to produced pauses than the short pauses. Furthermore, the perceived short pauses, which are not related to produced pauses, are considerably more than the perceived long pauses.

Figure 1c shows the total number of perceived themes by listeners and readers. The total number of perceived thematic units by listeners and readers is approximately the same. Most thematic units, on the other hand, by both listeners and readers, are related to produced pauses. The unrelated to produced pauses themes by readers are however more than that by listeners.

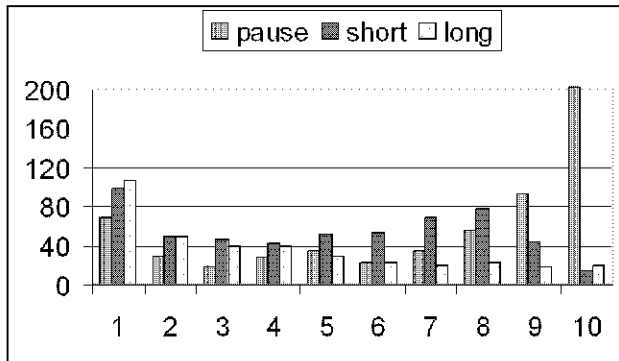
Figure 1d shows the total number of perceived hyperthemes and hypothemes by listeners and readers. The total number of hyperthemes by listeners and readers is approximately the same. More hyperthemes are however related to produced pauses than hypothemes. Hypotheses by listeners, on the other hand, are more than hypotheses by readers. More hypotheses than hyperthemes by listeners are unrelated to produced pauses. Hypotheses by listeners are also more in comparison to both hyperthemes and hypotheses by readers.



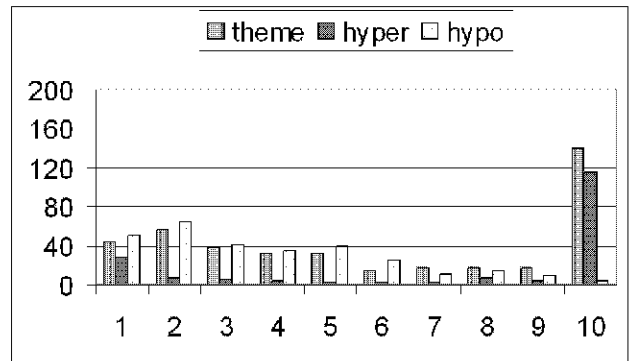
(a)



(c)



(b)



(d)

Figure 2. Pauses and themes frequency agreements (1-10): (a) total (pause) as well as short and long perception of produced pauses, (b) total (pause) as well as short and long pauses perceived, (c) total (themes) as well as hyperthemes (hyper) and hypothemes (hypo) by listeners, (d) total (themes) as well as hyperthemes (hyper) and hypothemes (hypo) by readers.

Figure 2a shows perception agreement of produced pauses among ten (1-10) listeners. Perceived short and long pauses as well as perceived pauses in total (irrespective short or long) are shown. The general pattern of perceived pauses in total is a right-headed one, according to which this agreement is overwhelmingly above any other perceptual agreement, represents almost half of the frequency of pauses (200 out of 444). The agreement of perceived pauses as either short or long ones is however rather evenly distributed.

Figure 2a shows agreement of perceived pauses among ten (1-10) listeners. Perceived short and long pauses as well as perceived pauses in total are shown. The general pattern of this Figure is much the same as the previous one with a right-headed pattern of agreement among the ten listeners. Minimal agreement among listeners (1) corresponds however to more perceived pauses in comparison to the previous figure, which is an indication that there is more agreement of perceived pauses in relation to produced ones.

Figure 2c shows agreement of perceived themes among ten (10) listeners. Perceived hyperthemes and hypothemes as well as perceived themes in total are shown. With reference to the total perception of themes, the general pattern of this Figure is much the same as the previous ones (2a, 2b), with a right-headed pattern of agreement among the ten listeners. The right-headed pattern of this Figure is also evident with reference to hyperthemes but not the hypothemes.

Figure 2d shows agreement of perceived themes among ten (10) readers. Perceived hyperthemes and hypothemes as well as perceived themes in total are shown. With reference to the total perception of themes, the general pattern of this Figure is much the same as the previous ones (2a, 2b, 2c), with a right-headed pattern of agreement among the ten readers. The right-headed pattern of this Figure, much like the previous Figure (2c) is also evident with reference to hyperthemes but not the hypothemes.

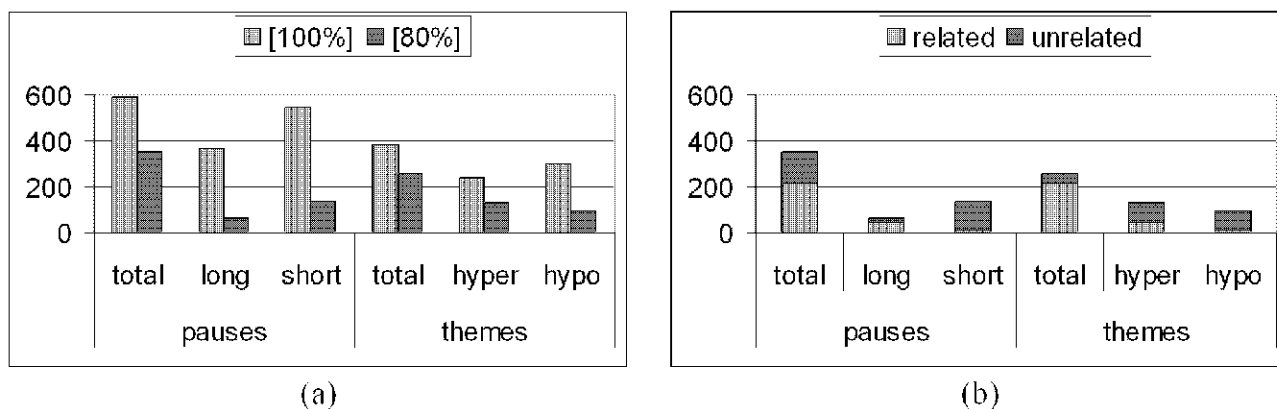


Figure 3. Perception of pause and thematic categories by listeners: (a) total number (100%) and 80% agreement among listeners and (b) pause-to-theme relations based on 80% calculations.

Figure 3a shows the perceived pauses and themes in total as well as at least 80% agreement among the ten listeners. The 80% agreement drops substantially with reference to all pause and theme categories.

Figure 3b shows pause-to-theme and theme-to-pause related and unrelated data. The number of pause-to-theme unrelated data is bigger than vice versa. Long pauses and hyper-themes, on the other hand, are more related than short pauses and hypothemes.

Discussion and conclusions

The main results of the present study indicate the following: First, more pauses are perceived than produced. Second, listeners and readers segment continuous speech and continuous text respectively in both hyperthematic and hypothematic units with considerable variability. Third, higher hyperthematic units have more segmentation constancy among both listeners and readers than lower hypothematic units.

Although we do have provided experimental evidence for thematic segmentation of speech and text, the large variability among both listeners and readers remains an open question. We assume that, in a speech communication process, a linguistic category, in principle, has both acoustic and perception correlates which are available to listener. Perception studies on stress and focus in Greek, which are related to lexical level and sentence level speech production, respectively, may come up to well over 90% identification (e.g. Botinis 1989; Botinis, Fourakis and Gawronska 1999, Botinis, Bannert and Tatham 2000). In less optimal contexts, such as postfocus ones, identification of stress distinctions may however drop to about 70% (Botinis 1989). One-minute radio

broadcast news is one of the most optimal types of material with reference to thematic segmentation and we would expect much more agreement of thematic identification, at least among listeners in the first place.

References

- Botinis A. (1989). *Stress and Prosodic Structure in Greek*. Lund University Press.
- Botinis A., Bakakou-Orphanou A., Tzimokas D. and Patsalou Th. (2004). Thematic segmentation and syntactic correlates. *Proceedings of the Prosody and Pragmatics Conference* (Preston, UK, forthcoming).
- Botinis A., Bannert R. and Tatham M. (2000). Contrastive tonal analysis of focus perception in Greek and Swedish. In Botinis, A. (ed.), *Intonation: Analysis, Modelling and Technology*, 97-116. Dordrecht: Kluwer Academic Publishers.
- Botinis A., Fourakis M. and Gawronska B. (1999). Focus identification in English, Greek and Swedish. *XIV International Congress of Phonetic Sciences*, vol. 2, 1557-60, San Francisco, USA.
- Botinis A., Gawronska B., Bannert R. and Sandblom E. (2003). Crosslinguistic analysis and modeling of prosody. *Proceedings of the XVth International Congress of Phonetic Sciences* (Barcelona, Spain), 1305-1308.
- Botinis A., Gawronska B., Katsika A. and Panagopoulou D. (2003). Prosodic speech production and thematic segmentation. *Proceedings of the Swedish Phonetics Conference FONETIK 2003, PHONUM 9*, 113-116.