

SEGMENTAL AND SUPRASEGMENTAL CUES TO PROSODIC BOUNDARIES

Evia Kainada
evia@ling.ed.ac.uk

Abstract

This paper addresses the issue of whether the application of different types of phonetic processes can be accounted for by the same prosodic structure. Most previous work in Laboratory Phonology has implicitly taken for granted that durational, intonational and external sandhi phenomena are all regulated by the same prosodic structure. No evidence as such for the specific assumption has been presented, though. In this paper I am presenting the assumptions made by previous research, some misconceptions in the field, and the importance of examining the particular topic. I am also presenting the reason for looking at several phonetic processes at the same time in order to examine whether they are organized by the same structure, and if so, whether they work at the same level or not.

1. Introduction

My research will examine whether segmental sandhi processes, durational processes, and intonational marking of constituents are all regulated by the same structure, and can thus be considered as cues to that structure. The reason for choosing these specific phenomena is that they have been traditionally considered as potential cues to a prosodic structure.

The main questions addressed are the following:

1. Can segmental and suprasegmental phenomena, like the ones just mentioned, be considered as cues to the same structure/hierarchy?
2. What is the nature of their relation in trying to identify this constituency?

Some following questions are:

3. If sandhi phenomena are considered to have a gradient output can they still signal prosodic constituents?
4. Do all phonetic phenomena operate on the same levels of a Prosodic Structure, or do some operate on higher and others on lower levels?
5. Is the manifestation of the durational (and intonational) phenomena gradient as well?

In this paper I will present different theories of Prosodic Phonology who have used phonetic cues to explain how we organise speech, and also how does Phonology relate to other components of the grammar. I will present some limitations of previous research, suggesting that further research is in order, as well as the motivation for my proposed research. I will be working with Modern Greek, with a subsidiary goal of examining how external sandhi phenomena work in this particular language. I will try to identify the exact 'units' of speech (or else categories, levels, or constituents) that each of the different processes define, as well as identify the Prosodic Structure of Greek.

2. Definitions

The investigation I propose focuses on processes that are regarded as cues to the prosodic structure of speech. The phonetic processes I will focus on are pre-boundary lengthening, external segmental sandhi, and intonational patterns. The use of those terms is not always consistent. Therefore, I will first present the way I will be using those terms.

With the term *pre-boundary lengthening* I refer to the lengthening effect that the prosodic boundary exercises on any part of the word preceding the boundary (see e.g. Wightman, Shattuck-Hufnagel, Ostendorf & Price 1992, Beckman & Edwards 1990 etc.).

By the term *intonational patterns* I refer to the way intonational contours have been found to group chunks of speech in a hierarchical manner (e.g. Pierrehumbert 1980, see Ladd 1996 for an introduction).

Finally, by saying *external segmental sandhi* I refer to the rules that describe the way segments occurring across words may change their pronunciation. They are purely phonological rules (Nespor & Vogel 1986) which take place at the post-lexical component of the phonology. For Modern Greek it has been postulated that one such phenomenon is ‘stop-voicing’ (Pelekanou & Arvaniti 2001 among others). External sandhi, segmental sandhi, or sandhi phenomena are thus the phenomena that take place between words and whose application is thought to depend on prosodic conditions.

Recently, a distinction between *categorical* and *gradient* sandhi has arisen¹. At this point I will explain my use of those terms.

i. Phonetic Gradience: the phonetic output of a sandhi phenomenon can vary. For example, the application of vowel hiatus in Greek has been found to have a variable phonetic output; from a total deletion of one of the two vowels, to the merging of the formants of the two vowels (resulting in one vowel), or the partial assimilation of two vowels, to the full existence of both vowels (e.g. Baltazani 2006). The phonetic output of a sandhi phenomenon can be influenced by several different prosodic factors, such as stress etc.

ii. Statistical Gradience: this type of gradience refers to the probability of a sandhi phenomenon to occur across prosodic constituents. For example, the application of vowel hiatus in Greek is more likely to occur within the domain of prosodic words, than across two intermediate phrases, but the latter is still a possibility.

These two types of gradience have been mixed in the literature, while I would like to keep them separated.

In the following section several theories that refer to the way the phonological component relates to other components of the grammar will be presented, as well as how these theories have been using phonetic phenomena to support their proposals.

3. Prosodic Phonology

Prosodic phonology is a theory that tries to explain the relationship between phonology and the other components of the grammar, mainly the syntax. Its main claim is that there is a prosodic structure, which is divided into prosodic domains, which in turn are hierarchically organised. This structure is similar to the way the syntactic trees look like, but with many different constraints on its formation. It is thought to be in close relationship with, but not isomorphic to, syntax (for a full account of this theory, its claims etc. see Selkirk 1978, 1981, 1984 and 1986, Nespor & Vogel 1982 and 1986, and Hayes 1989).

For example, although the prosodic structure of an utterance is often similar to the syntactic one, there are times when the prosodic structure (signalled by pauses, intonational

¹ This is a tentative set of names that we use. I am grateful to Prof. D. R. Ladd for suggesting them.

contours, etc.) differs significantly from the syntactic one. A typical example for the discrepancy is the following:

A. [This [is [the cat that killed [the rat that ate [the malt]_{NP}]_{NP}]_{VP}]_S

B. [This is the cat] [that killed the rat] [that ate the malt]

Sentence A shows the syntactic configuration, while sentence B the actual way of producing the sentence in chunks.

There were several proposals that tried to account for this discrepancy. The most common one was that phonological information was formed after syntactic organisation was ready in a linear way. A different proposal was presented by Halliday (1967), who suggested that the formation of intonational groups was based on ‘information structure’ (based on intonational contours he had identified). According to Gee & Grosjean (1983), there is a prosodic structure based on the occurrence of pauses and final lengthening.

One of the most influential proposals was the one presented by Selkirk (1978), according to which there is an autonomous system, namely the prosodic/intonational structure, which is separate from the syntactic structure. This structure is made up by ‘chunks’, which are often called prosodic domains (or constituents, or levels) of the prosodic structure. Selkirk (1981) defines the prosodic structure and its levels as follows:

“The phonological units [...] will be referred to as prosodic categories and the hierarchically arranged suprasegmental portion of phonological representation will be referred to as prosodic structure” (p. 381).

By that time it was obvious that the relationship between the prosodic constituents and the syntactic constituents was not straightforward.

There were also several discussions about the access that phonology had to syntax, and whether it is syntactic information that the prosodic structure takes into consideration. These proposals were the ones that set out the phonology-syntax interface. One proposal suggested a direct access of phonology to syntax, according to which each type of phonological rule would occur at different times with reference to the construction of the syntactic tree (Kaisse 1985).

The other approach was the indirect access of phonology to syntax, which took place via the mediation of the prosodic structure. The prosodic structure is different to the syntactic one, but its construction needs to refer to it. The idea of syntactic knowledge being accessible to phonology rose the issue of what types of syntactic information was the prosodic structure paying attention to. The two main proposals that were put forward were the End-Based approach by Nespor & Vogel (1986) and the Relation-Based approach by Selkirk (1984).

Recently another approach has been proposed by Steedman (1991), who suggested that the phonetic output of our utterances is organised by what is called the Combinatory Categorical Grammar, which has access to semantic information (which in its turn is very important also for the construction of syntactic trees), and which groups intonational contours together.

These last three approaches have remained up to our days as some of the most influential approaches to the analysis of the prosodic structure of speech. They have proposed particular levels of a prosodic structure, which are still widely used in the field of laboratory phonology. In order for these approaches to present evidence for the existence of the

Prosodic Structure they were postulating, they used particular phonetic phenomena occurring in the languages they were investigating. The most commonly used phonetic cues in the literature are external sandhi and intonational contours, alongside with some phonotactics. This is where our research relates to theirs, since it will mainly examine whether these phonetic cues are organised by the same structure.

4. Identification of prosodic levels, limitations and motivation

4.1. Use of phonetic cues

The main phonetic cues used by researchers when identifying prosodic constituents are durational patterns at the edges of prosodic domains, intonational contours, external sandhi, and sometimes phonotactics. Most of the researchers mentioned above made use of such phonetic processes to validate their proposal. Nespor and Vogel (1986) for example, used principles of syllabification for the prosodic level of the *syllable*, segmental phonological rules (such as aspiration in English), autosegmental association rules, and phonotactics for the prosodic level of the *foot*, external sandhi such as nasal deletion for the domain of the *prosodic* word etc. Steedman (1991) referred to intonational contours, and some external sandhi, while Selkirk (1984) referred to durational patterns, intonational patterns, external sandhi, and phonotactics, among others. Gee & Grosjean (1983) mainly referred to durational patterns and the insertion of pauses and final lengthening. Finally, Cooper & Sorensen (1981) referred to intonational contours to identify what they call the ‘speaker’s mental computations’.

All the above show that there are several phonetic processes which are important for the identification of how different components of the grammar work together. There has been no investigation, to our knowledge, examining several phonetic cues at the same time in order to test whether they can all be regarded as cues to the same structure. Thus, the assumption that different types of cues, such as segmental and suprasegmental, can be considered as cues to the same structure needs to be revisited and the theories based on that assumption may need to be re-examined.

4.2. Limitations of the existing studies

Previous studies attempting to address the issue of how we signal different levels of a potential prosodic structure suffer from one or more of the following issues: a) they are based on the *assumption* that different types of phenomena work under the same structure, b) the analyses are either incomplete because they use impressionistic data or they are based on *misconceptions*, and c) when referring to connected speech processes, such as external sandhi, they assume that these processes are *categorical*.

Assumption

The first point of the limitations of the previous research involves the assumption regarding the application of phonetic cues. This assumption can be found often in often in the literature in cases like the following taken from (Selkirk 1981: 385):

“The claim is that suprasegmental phenomena such as intonation, phrase stress and rhythm, and segmental phenomena such as external sandhi, rely crucially on the parsing of the utterance into these phrase types”.

The same assumption is also found in more recent investigations. For example, Frota (1998) mentions:

“...various types of evidence, besides sandhi, may be considered in the identification and definition of prosodic domains” (p. 4).

The fact remains that this assumption has not been tested as such.

Misconceptions

One further limitation of the previous research in identifying the Prosodic Structure has been the fact that in many cases researchers were based on impressionistic data. On several occasions we can see that analyses by researchers such as Nespor and Vogel don't hold under scrutiny. For example, their analysis of *nasal deletion* suffers when placed under close inspection by native speakers. As a native speaker of Greek I would have to disagree with the examples provided, e.g. when mentioning that nasal deletion never takes place between the two following words in Greek:

πριν φάω
 /'prin 'fao/
 ['prin]_c['fao]_c → *['pri'fao]
 (=before I eat)

It may be the case that the particular sequence with the nasal deleted does not sound as natural, but it still occurs in several occasions. This could be a case of a sandhi phenomenon with a statistically gradient output. It is not clear what the probability would be for this phenomenon to occur in each of the proposed prosodic constituents, but it looks plausible that there would be a low probability of it occurring in the level that I am investigating. Similar problems with impressionistic data can be found in the analyses of Greek sandhi by Kaisse (1985) and Condoravdi (1990) for vowel hiatus.

A different kind of problem is that different languages have been used for the identification of different levels of the Prosodic Structure. This way it becomes difficult for the researcher to be certain whether all the proposed levels would actually stand in one language. I propose that one language is investigated in its entirety, so that safer conclusions can be drawn. This is the reason why I restrain my research in one language, namely Modern Greek.

Gradient vs. Categorical

As mentioned before, Nespor and Vogel (1986) did not take into consideration the possibility of the phonetic output of phonetic processes being gradient rather than categorical. This is an oversight that makes us extremely cautious as far as their generalisations regarding Prosodic Levels are concerned. For example, they investigated the application of *s-voicing* in Modern Greek. They supported that it only occurs within intonational phrases and not across them. This proposal was tested by subsequent research and was found to be incorrect. In fact, Pelekanou & Arvaniti (2001) investigated this phenomenon (amongst other segmental sandhi phenomena of Greek) in naturally occurring speech, and they found that it is not a phonological rule and that it can occur across any level of the prosodic structure.

4.3. Motivation for the current research

Apart from the limitations of previous studies, there are also further reasons that have motivated my research questions. I will report on two examples taken from Greek, suggesting that the application of phonetic processes and their correlation at identifying prosodic constituents is still a matter of further debate.

1. There are cases that do not abide to the normal organisation of constituents, as proposed by Arvaniti & Baltazani (2005). For example, Arvaniti (1999) reports on a case of stop voicing, where stop voicing doesn't occur when we would expect it

to. Stop-voicing in Greek is considered to occur within the boundaries of a Prosodic Word. In the following example, though, this is not what we find:
 των κήπων τους → /ton (=article-genitive plural) 'kipon (=noun) tus (=clitic)/ (= of their gardens) → [to'gipontus]

In this case, stop voicing occurs between the article and the noun, but not between the noun and the clitic. This is interesting, given that they both theoretically belong to the same prosodic constituent. It might be the case that the sandhi phenomenon depends on different types of information, other than the prosodic boundaries. This is one of the vital questions that our research will try to address.

2. It can be the case that different phonetic processes apply on different levels of the prosodic structure. Kainada (2005) investigated the application of durational patterns and intonational contours in Greek in prosodic words of the sort:

το μέλι μου

[to 'meli mu] = my honey (one prosodic word)

[to 'meli] [mu...] = the honey... (prosodic boundary between /'meli/ and /mu/).

It was found that there was no difference in the intonational contour used, meaning that it might be the case that lower constituents in Greek (even of the type of prosodic word) are not signalled by intonational contours.

5. Conclusion

5.1. Research Questions - Summary

Prosodic structure has been found to differ from the syntactic one. Therefore, researchers have been trying to find the organisation of this structure, as well as the way it relates to other components of the grammar, such as the syntactic component. Some suggest that the way we organise the phonological structure is by referring to syntactic information, while others say that it is semantic information that phonology relates to.

No matter which theory one wants to subscribe to, evidence to support each of these theories is needed. This evidence is based on phonetic phenomena and different theories make use of the same phonetic phenomena to support their proposals. The main issues that arise from the research presented so far in the field are:

- Although all researchers seem to use the same phonetic processes to support their proposals, there has been no research so far, to our knowledge that investigates the application of different types of cues at the same time.
- Different types of phonetic cues have indeed been identified as procedures that seem to operate in a hierarchical way. We do not know, though, if all different cues are organised by such a hierarchy, whether they would demarcate the same hierarchy, and whether they would operate on the same levels of that hierarchy.
- It is also not clear yet whether phonetic phenomena, such as external sandhi and durational patterns behave in a gradient of categorical manner.

5.2. Proposed research

Given the problems mentioned so far, my research will examine the following:

- The prosodic structure of one language, i.e. Standard Modern Greek, so as to have a complete picture of how different phonetic processes operate. I will be investigating my native language, i.e. Standard Modern Greek.

- Three phonetic cues at the same time, i.e. pre-boundary lengthening, intonational contours and segmental sandhi, so as to see whether they can be predicted by the same structure.

The following research questions will be addressed:

- Can intonational marking, durational lengthening and sandhi processes in Greek all be predicted by the same structure?
- How is each of those processes manifested in Greek?
- Are those processes gradient or categorical? How many levels can they distinguish?
- Do all these processes operate on the same level of the prosodic structure, or is it possible for some of them to operate on a higher level, while others on a lower?

It has become clear in this paper that our main concern is whether the cues that have been presented for the prosodic hierarchy are all reliable to be thought as cues for the same hierarchy. We have seen some cases which could suggest otherwise, but what we mainly have not seen is evidence to convince us that this is really the case.

It has been an implicit (and sometimes explicit) assumption of many studies, that these cues point to the same structure, and that thus different domains arising from different cues can actually be equated.

Answers to which are the cues that signal the prosodic domains and how that happens are very important not only for the theory of the prosodic structure itself and our understanding of the phonological organisation of the speech, but also for the modelling of speech production and the generation of fully functioning tools for speech synthesis and speech recognition.

5.3. Implications for existing research

If my research shows that the application of the three phenomena is organised by the same structure, then all the theories mentioned so far will be justified to use them. However, if the results are not clear-cut the implication for those theories will be of great importance. Either they would need to reconsider the evidence for some of the constituents, or they would need to reconsider the whole use of some of those phenomena, depending always on the findings.

Furthermore, the investigation of the external sandhi phenomena aims to answer important questions regarding the phonetics/phonology interface. It will be of interest to see whether the gradient output of sandhi phenomena can fit in a theory of prosodic domains, and if so, whether these are the same domains that are signalled by lengthening and intonational phenomena. These results will be important not only for theories of a Prosodic Structure, but also for Articulatory Phonology and initial strengthening, given that this it will be shown whether there could be a correlation between the application of different phenomena.

Finally, my research will, among other things, investigate the phenomenon of stop voicing and pre-boundary lengthening in Standard Modern Greek, which have not been a matter of investigation in the past.

6. Proposed Study

6.1. Phenomena to be tested

As mentioned previously, my research will focus on three processes, namely, *pre-boundary lengthening*, *intonational patterns* and *external sandhi*. For external sandhi I will mainly be investigating the application of *stop voicing* (and potentially *vowel hiatus*), given that it has not been examined in the past for Greek, and it shows quite an irregular application.

6.2. Sample Materials

In an attempt to investigate different processes at the same time, materials that will be able to accommodate this are in order. One case of potentially interesting materials that could be used, should I decide to use elicited over spontaneous speech is the following:

[κοίτα με] [συνέχεια]
/'kita me si'neçia/ = [look at me] [constantly]

[κοίτα] [με το κιάλι]
/'kita me to ci'ali/ = [look] [with the binoculars]

[κοίτα], [με το κιάλι σταθερό]
/'kita me to ci'ali staðe'ro/ = [look] [with steady binoculars]

In those examples the duration of /ta/ in /'kita/ is being examined for pre-boundary lengthening. Similar sentences with nasals would have to be built in order to get intonational measurements at the same positions and examine whether constituents are also (assuming that they are durationally) intonationally marked.

6.3. Tasks

A difficult decision that I will have to take is whether to use spontaneous or laboratory speech in my production experiments. One possibility is to use a Map Task and direct my speakers to produce the required segments by manipulating the names of the landmarks. Should that idea fail, there is also the possibility to use cartoons and ask my participants to describe what they see, while being able to intervene and try to elicit the required sentences by them². In order to get a first idea of the results I am planning to conduct a pre-pilot experiment investigating laboratory directed speech.

6.4. Hypotheses

There are three competing hypotheses that arise from my research questions:

- There is one structure overarching all processes, meaning that all phonetic processes can be considered as cues to the same structure.
- There is one structure, but different processes refer to different levels of this structure.
- There are different structures organising different processes, such as segmental and suprasegmental processes.

Any of those three outcomes would have important implications for the research being conducted as we speak, for theories regarding Syntax-Phonology mapping and for the Phonetics-Phonology interface.

² These ideas were kindly suggested to me by my supervisors, Prof. D. R. Ladd and Dr. A. Turk.

References

- Arvaniti, A. (1999). 'Greek voiced stops: prosody, syllabification, underlying representations or selection of the optimal?', in *Proceedings of the 3rd International Linguistics Conference for the Greek Language*. Athens: Ellinika Grammata, pp. 383–390.
- Arvaniti, A. & M. Baltazani (2005). 'Intonational analysis and prosodic annotation of Greek spoken corpora', in S.-A. Jun (ed.), *Prosodic Typology: The Phonology of Intonation and Phrasing*. Oxford: Oxford University Press, pp. 84–117.
- Baltazani, M. (2006). 'Focusing, prosodic phrasing, and hiatus resolution in Greek.' To be published in *Papers in Laboratory Phonology VIII*.
- Beckman, M. E. & J. Edwards (1990). 'Lengthenings and shortenings and the nature of prosodic constituency', in J. Kingston & M. E. Beckman (eds.), *Papers in Laboratory Phonology I: Between the Grammar and the Physics of Speech*, Cambridge: Cambridge University Press, pp. 152–178.
- Condoravdi, C. (1990). 'Sandhi rules of Greek and prosodic theory', in S. Inkelas & D. Zec (eds.), *The Phonology-Syntax Connection*. Chicago: University of Chicago Press, pp. 63–84.
- Cooper, W. E. & J.M. Sorensen (1981). *Fundamental Frequency in Sentence production*, New York: Springer.
- Fallon, P. (1993). 'Naturally occurring hiatus in Modern Greek', in I. Philippaki-Warbuton, K. Nicolaidis & M. Sifianou (eds.), *Themes in Greek Linguistics, Papers from the First International Conference on Greek Linguistics*. London: Benjamins, pp. 217–224.
- Frota, S. (1998). 'Prosody and Focus in European Portuguese.' PhD thesis, Universidade de Lisboa.
- Gee, G. P. & F. Grosjean (1983). 'Performance structures: A psycholinguistic and linguistic appraisal.' *Cognitive Psychology* 15, pp. 411–458.
- Halliday, M. A. K. (1967). *Intonation and Grammar in British English*. The Hague: Mouton.
- Hayes, B. (1989). 'The prosodic hierarchy in meter', in P. Kiparsky & G. Youmans (eds.), *Phonetics and Phonology 1: Rhythm and Meter*. San Diego: Academic Press, pp. 201–260.
- Kaisse, E. M. (1985). *Connected Speech: The Interaction of Syntax and Phonology*. Florida: Academic Press.
- Kainada, E. (2005). 'Interaction of Cues to Phrasing and Cues to Lexical Stress in Greek.' unpublished MSc thesis, University of Edinburgh, Department of Theoretical and Applied Linguistics.
- Ladd, D. R. (1996). *Intonational Phonology*. Cambridge: Cambridge University Press.
- Nespor, M. & I. Vogel (1982). 'Prosodic domains of external sandhi rules', in H. van der Hulst & N. Smith (eds.), *The Structure of Phonological Representations - Part 1*. Dordrecht: Foris, pp. 225–255.
- Nespor, M. & I. Vogel (1986). *Prosodic Phonology*. Dordrecht: Foris.
- Pelekanou, T. & A. Arvaniti (2001). 'Postlexical rules and gestural overlap in a Greek spoken corpus', in *Proceedings of the 5th International Conference on Greek Linguistics*.
- Pierrehumbert, J. B. (1980). 'The Phonology and Phonetics of English Intonation.' PhD thesis, Massachusetts Institute of Technology.
- Selkirk, E. (1978). 'On prosodic structure and its relation to syntactic structure', in T. Fretheim (ed.), *Nordic Prosody 2*. Trondheim: TAPIR, pp. 111–140.
- Selkirk, E. (1981). 'On the nature of phonological representation', in T. Mayers, J. Laver & J. Anderson (eds.), *The Cognitive Representation of Speech*. Amsterdam: North Holland, pp. 379–389.
- Selkirk, E. (1984). *Phonology and Syntax: the Relation Between Sound and Structure*. Cambridge, MA: MIT Press.

- Steedman, M. (1991). 'Surface structure, intonation, and "focus"', in E. Klein & F. Veltman (eds.), *Natural Language and Speech, Proceedings of the ESPRIT Symposium*, Brussels, pp. 260–296.
- Wightman, C. W., S. Shattuck-Hufnagel, M. Ostendorf & M.J. Price (1992). 'Segmental durations in the vicinity of prosodic phrase boundaries.' *Journal of the Acoustical Society of America* **91**, pp. 1707–1717.