

The phonetic and phonological nature of prosodic boundaries: evidence from Modern Greek

Evia Kainada



A thesis submitted in fulfilment of requirements for the degree of
Doctor of Philosophy

to

Linguistics and English Language
School of Philosophy, Psychology and Language Sciences
University of Edinburgh

2009

Declaration

I hereby declare that this thesis is of my own composition, and that it contains no material previously submitted for the award of any other degree. The work reported in this thesis has been executed by myself, except where due acknowledgment is made in the text.

Evia Kainada

Abstract

Research on prosodic structure, the underlying structure organising the prosodic grouping of spoken utterances, has shown that it consists of hierarchically organised prosodic constituents. The present thesis explores the nature of this constituency, in particular the question of whether prosodic structure is comprised of a given set of qualitatively distinct domains, or of a set of domains of the same type varying only gradiently in “strength”, or a possible mixture of both types of relations across prosodic levels.

This question is addressed by testing how prosodic constituency (mirrored on boundary strength manipulations) is signalled acoustically via pre- and post-boundary durations, intonation contours, and two sandhi processes, namely vowel hiatus resolution and post-nasal stop voicing in Modern Greek.

Results show that the phonetic signalling of boundary strength provides support for a mixture of both differences of *type* and *strength* across prosodic levels, with some levels only differing in terms of their strength. Pre-boundary durations and resolution of vowel hiatus are gradiently affected by boundary strength with shorter to longer durations from lower to higher domains, and less instances of vowel deletion higher in the hierarchy. Post-nasal stop voicing is qualitatively affected by boundary strength with almost all voicing instances occurring in the lowest constituent of the structure in the way a qualitative view of prosodic constituency would predict, and in line with research on prosodic phonology. Finally, both the alignment and scaling of intonation contours at the edges of domains is found to distinguish qualitatively the lowest domain from the higher ones. All higher phrasal domains align with respect to the boundary and their peak scaling varies consistently gradiently

across speakers. When combining those two findings, support is provided for the existence of differences of strength and type within the same process.

Taken together the results from these four phenomena support the postulation of an underlying prosodic structure with a limited number of qualitatively distinct domains, within which at the same time some type of recursivity or structured variability must be allowed for. It is shown that there are structural properties of speech, like the length of the utterance, influencing the organisation of utterances in a principled gradient manner, supporting the existence of differences of strength within domain types. These findings bear significance for theories of prosodic structure that have assumed either the view of solely qualitative differences, or sole boundary strength differences, as well as for future proposals on prosodic constituency.

Finally, the use of Modern Greek in this thesis adds to the existing literature on a language that has been extensively used by researchers working in views supporting the existence of qualitative distinctions of type across prosodic domains, and provides the first in depth experimental analysis of post-nasal stop voicing.

Acknowledgements

First of all, I would like to thank my two supervisors, Bob Ladd and Alice Turk. I have been extremely lucky to work with them, and I can't imagine starting or finishing this PhD without their constant help, support, encouragement, and patience. I need to express my deepest gratitude to them for helping me since before the beginning of the PhD. Bob's influence has been vital in my orientation, and the development of the project from the first to the last day. He has been a constant support and a person urging me to always do more. Alice has spent endless hours talking about everything from the experimental design to general theoretical concepts. Her comments and input have always been challenging, helping me move forwards. Their ideas and way of thinking have been a great influence to me. Of course, I will not only miss the professional side of our communication, but also the personal mentoring and supporting role they both have had since my MSc.

Before moving on to expressing thanks to individuals, I would like to express a more general 'thank you' to the Department of Linguistics and English Language, to the Center for Speech Technology Research, to the School of Informatics, and to the people that make all those departments such a brilliant place to be, and such a vibrant and interesting community. I have had so many opportunities while being in Edinburgh, that I wouldn't have hoped for. I have experienced what interdisciplinary research really is all about, joined a number of research groups, worked in interesting projects, organized seminars and meetings, and was given the opportunity to contribute actively to the academic and social life of all those places.

During my work on this thesis, I have received a lot of help, support and guidance from a number of people. I need to thank Satsuki Nakai and Sasha Calhoun for spending so much time talking to me at various points during this thesis. Mits Ota for his insightful comments during my postgraduate and P-Workshop talks, and Bert Remijsen, who has always been very generous with his knowledge of recording techniques, scripting, and statistics. Marianne Pouplier for her feedback on my work when she was still in Edinburgh. Simon King was extremely helpful in so many parts of my life in Edinburgh; apart from being there for my MSc studies, he offered me work on several projects, helped with funding for PhD and post-doc applications, and offered his support and guidance without holding back.

On the technical side of things, the computing support of the Department of Linguistics has made my life a lot easier; Eddie Dubourg, Mike Bennett, and Ziggy Cambell from Linguistics, and Mike Allerhand from Psychology, have offered their help bounteously.

I would also like to thank people that I met along the way of the PhD. First, I would like to thank Cécile Fougeron, who welcomed me in her Department at the early stages of this thesis, and gave me advice on many aspects of the thesis. Jean Carletta for help with statistical analysis. Mary Baltazani has been a significant help with GrToBI issues. I would also like to thank Amalia Arvaniti for her support during the final stages of the thesis.

On a more personal level, I have been extremely lucky to make very good friends in Edinburgh, who have made my life full and happy, even at times when the PhD could have been very lonely. Peter Bell, Catherine Dickie, Anna Leonard-Cook, Kate Messenger, Jeff Mitchell, Emi Sakamoto, Silke Scheible and Frances Wilson have been my pillars of support. Peter has not only been a friend, but has also been a brilliant LaTeX-buddy, to whom I owe very much gratitude for helping me in really tricky situations. Catherine has been the best office-mate ever, a great friend and a brilliant phonetics/phonology resource. Silke, Frances and Kate have offered me their unconditional friendship, and I will miss more than anything our lunch and coffee breaks in the future.

Finally, I would like to thank various people that have been in my life the past few years and have made it possible to finish the PhD and feeling like it was a very enjoyable experience. Frank Keller and Colin Matheson, Christian Geng, Tareq Maiteq, Tim Mills, Cassie Mayo, Rob Clark, Joanna Keating, Dan Wedgwood, Lynda Hutchinson, and all the PhD students of the Department of Linguistics and English Language. Also, I would like to thank the many Greek friends I made in Edinburgh, especially Manolis Stavrou, Georgia Vendra, Stavros Parlalis, and Georgios Petkos, and in general the Hellenic Society for organising Greek dancing classes, and for being a lively community with many activities.

I am indebted to the Economic and Social Research Council for financially supporting this research, as well as to the Department of Linguistics and English Language and to the University of Edinburgh for various travel grants.

These past few years would have been unbearable if it weren't for the love and constant support I received from Angelos Lengeris. I would like to thank him for believing in me, standing by me, and for being a constant source of inspiration. Finally, I would also like to thank my brother, Dimitris Kainadas for his continuing support, his humour during bad times, and his advice at all times.

Last, but definitely not least, I would like to thank my parents, Elias and Polina, for supporting me in any way possible. I cannot begin to thank them for giving me options, urging me to always be better, and making sure that I have everything needed to achieve my goals. I am certain that I would not have made it without them. Thank you very much for always being there for me.

Contents

Declaration	i
Abstract	ii
Acknowledgements	iv
Chapter 1 Introduction	1
1.1 Prosody and Boundary Strength	2
1.1.1 Direct syntax-prosody mapping	3
1.1.2 Prosodic structure as consisting solely of differences of bound- ary strength	6
1.1.3 Prosodic structure as consisting of domains of different type	9
1.2 Do labelled domains exist?	13
1.2.1 Evidence against labelled domains: from sandhi	13
1.2.2 Evidence against labelled domains: from intonation	19
1.2.3 Terminology and acoustic evidence for domain types	26
1.2.3.1 Nature of process investigated	27
1.2.3.2 Types of boundary strength effects	29
1.2.3.3 Statistical findings in support of different domain types	31
1.3 This Thesis	33
1.3.1 Research Question	33
1.3.2 Methodological strategy	33
1.4 Summary and Overview of thesis	36
Chapter 2 General Experimental Design	38
2.1 General Design	38

2.1.1	Acoustic Correlates	39
2.1.2	Boundary conditions	41
2.1.3	Stimulus design	48
2.2	Method	50
2.2.1	Participants	50
2.2.2	Procedure and recordings	51
2.2.3	Analysis	53
2.3	Summary	54
Chapter 3 Pre- and post-boundary durations		55
3.1	Introduction	55
3.1.1	Durations as correlates to prosodic structure	55
3.1.1.1	Pre-boundary durations	56
3.1.1.2	Post-boundary durations	57
3.1.2	Durations and prosodic constituency identification in labelled views	58
3.1.3	Boundary-related duration patterns in Modern Greek	61
3.1.4	How far does pre-boundary lengthening spread?	63
3.2	Goals	64
3.3	Methodology	65
3.3.1	Testing for rate effects	65
3.3.2	Materials	65
3.3.2.1	Conditions BC1-BC5	65
3.3.2.2	Condition BC0	66
3.3.3	Segmentation	67
3.3.4	Measurements	68
3.3.5	Statistics	69
3.4	Results	70
3.4.1	Pre-boundary durations	70
3.4.1.1	Conditions BC1-BC5	70
3.4.1.2	By speaker analysis	75
3.4.1.3	Conditions BC0-BC5	77

3.4.2	Post-boundary durations	78
3.5	Conclusions	80
Chapter 4	Post-nasal Stop Voicing	84
4.1	Introduction	86
4.1.1	Definition of Post-Nasal Stop Voicing in Modern Greek	86
4.1.2	Phonology and Phonetics of voiceless and voiced stops in Modern Greek	88
4.1.3	Prosodic structure and post-nasal stop voicing in Modern Greek	90
4.1.4	Prosodic structure influences on stop consonant production	92
4.1.5	Research Goals and Chapter outline	93
4.2	Methodology	95
4.2.1	Production experiment	95
4.2.1.1	Materials	95
4.2.1.2	Segmentation	98
4.2.1.3	Measurements	98
4.2.2	Preliminary perception experiment	102
4.2.2.1	Design	102
4.2.2.2	Measuring inter-transcriber agreement	103
4.2.2.3	Analyses	105
4.3	Results	105
4.3.1	Production experiment on voiceless and voiced stops	105
4.3.1.1	Acoustic correlates of voicing distinctions	106
4.3.1.2	Boundary strength effect on voiceless and voiced stops	107
4.3.2	Preliminary perception experiment	110
4.3.2.1	Inter-transcriber agreement	111
4.3.2.2	Prosodic Boundary Effect	112
4.3.2.3	Lexical differences in the lowest domain	114
4.3.2.4	Conclusions on perception test	115
4.3.3	Production experiment of ‘nasal+voiceless stop’ sequences	116
4.3.3.1	Sandhi or coarticulation?	116
4.3.3.2	Qualitative or variable boundary strength effect?	121

4.3.3.3	Deletion of pre-boundary nasal	124
4.4	Conclusions	125
Chapter 5	Vowel Hiatus resolution	130
5.1	Introduction	131
5.1.1	Definition and nature	131
5.1.2	Vowel hiatus resolution and prosodic domain identification .	135
5.2	Goals and Overview of Chapter	139
5.3	Methodology	140
5.3.1	Production experiment	140
5.3.1.1	Materials	140
5.3.1.2	Segmentation	143
5.3.1.3	Measurements	145
5.3.2	Perception experiment	146
5.3.2.1	Procedure	146
5.3.2.2	A note on the procedure and terminology	147
5.4	Results	148
5.4.1	Preliminary perception experiment	148
5.4.1.1	Inter-transcriber agreement	148
5.4.1.2	Boundary strength effect	150
5.4.1.3	Conclusions on perception experiment	152
5.4.2	Production experiment	153
5.4.2.1	Sandhi or coarticulation?	153
5.4.2.2	Qualitative or variable boundary strength effect? .	157
5.5	Conclusions	162
Chapter 6	F0 alignment and scaling	166
6.1	Introduction	166
6.1.1	Intonation contours, prosodic constituents and gradience in intonation	166
6.1.2	Modern Greek and GrToBI	172
6.1.3	Effect of utterance length on F0 scaling	175
6.1.4	Acoustic support for labelled views	175

6.2	Goals	176
6.3	Expectations	177
6.4	Methodology	179
6.4.1	Materials	179
6.4.2	Measurements	180
6.4.3	Contour extraction and Normalisation Technique	180
6.4.3.1	Identification of turning points and smoothing of F0	181
6.4.3.2	Normalisation across speakers	182
6.4.3.3	Alignment	184
6.5	Results	184
6.5.1	Scaling	184
6.5.2	Alignment	188
6.5.3	Scaling and pre-boundary durations	190
6.6	Discussion	190
Chapter 7	Discussion	195
7.1	General Summary	195
7.2	Methodological contribution	197
7.3	Summary of chapters	198
7.4	Labelled domains and recursivity	202
7.5	Length effects	206
7.6	Concluding Remarks	207
Appendix A	Materials	209
Appendix B	F0	219
Appendix C	Stop Voicing	220
References	222

List of Tables

2.1	Exact numbers of sentences recorded for all speakers for each of the processes.	53
3.1	The eight items constructed for investigating pre- and post-boundary durations. Each of those was later embedded in five different prosodic conditions. Neuter nouns in Greek have the same morphological ending in the nominative and accusative plural. The same is true for many feminine nouns in the singular. The nouns used in these experiment had this feature, which means that they could be embedded in prepositional phrases using their accusative forms.	66
3.2	The six items constructed for investigating pre- and post-boundary durations in BC0. Each of those is based on the items constructed for conditions BC1 to BC5.	67
3.3	Mean duration (ms) of the onset, rime and whole syllable lengths of the ultimate, penultimate, antepenultimate and post-boundary syllables, along with the percent lengthening of BC2, BC3, BC4 and BC5 compared to BC1 on all pre-boundary syllables.	72
3.4	Mean duration (ms) of the pre-boundary rime for each of the six prosodic conditions including BC0 for only the two male speakers that were recorded producing the items.	78
4.1	Table showing the mean values of closure and VOT of all Greek voiceless stops, as presented in previous research, with results from this thesis added. Table adjusted from Arvaniti (2008). Results for voiced stops only come from this study, and what is marked as VOT refers to the duration of the release, and not the voice onset, since voiced stops in Modern Greek are fully prevoiced.	89
4.2	The 8 items constructed for investigating SV.	96
4.3	The 15 items constructed for measuring stop voicing in the lowest condition (BC0).	97
4.4	The 3 items constructed using voiceless and voiced stops not preceded by a nasal for baseline measurements. These items were embedded in the same conditions as their voiceless counterparts. The items for measuring voiceless stops were taken from the design for pre-boundary lengthening.	97

4.5	The number of items constructed for measuring stop voicing (SV), and the number of items for baselines, broken down by number of speakers, items, boundary conditions, repetitions, and the number of particles used.	98
4.6	Mean duration (ms) and amplitude (dB) of the closure and the release of phonologically voiceless and voiced stops, as produced by the 5 speakers.	106
4.7	Mean duration (ms) and amplitude (dB) of the closure and the release of phonologically voiceless and voiced stops, as produced by the 5 speakers for the pilot measurements.	108
4.8	Results in percentages for each rater. Rows represent the percentage that each segment was classified as a voiced or voiceless stop. Some small percentages might not be adding up to 100% given that the option of ‘No Answer’ was offered to the raters. Numbers in bold indicate instances of voiceless stops classified as voiced.	112
4.9	Results in percentages for classification of voiceless stops as voiced by all raters for each boundary condition. The percentage for the first boundary condition is higher, suggesting that stop-voicing mainly occurs on this level.	113
4.10	Differences between lexical items as to how many were classified as voiced, voiceless, or didn’t reach agreement.	114
4.11	Mean duration (ms) and amplitude (dB) values for all ‘nasal+voiceless stop’ sequences classified as voiceless and voiced by raters.	117
4.12	Mean values for all dependent variables in all prosodic conditions. The last column shows whether the effect of prosodic boundary strength was significant for each variable.	123
4.13	Percentage of nasals being classified as deleted or not across boundary conditions.	125
5.1	The six items constructed for each of the vowel pairs under investigation.	142
5.2	The six items constructed for each of the vowel pairs for BC0.	143
5.3	Mean percent of how many times [ao] and [oa] vowel pairs were classified by the raters as themselves, or as either of the baselines pooled across prosodic conditions.	150
5.4	Table showing the percentage values of how many times the vowel pairs [ao] and [oa] are classified as themselves or as one of the baselines for each prosodic condition across speakers (n=approx. 60 for each boundary condition for each vowel pair). The NA columns stands for the percentage rate that the raters did not reach an agreement (No Agreement).	151

5.5	Mean ERB values for all vowel pairs in all dependent variables, pooled across boundary conditions. Comparisons across vowel pairs are shown with Bonferroni post-hoc tests in the last row.	154
5.6	Mean ERB values from all [ao] and [oa] vowel pairs in all prosodic conditions. The effect of boundary strength was not significant for all any measurement.	159
5.7	Mean duration (sec) of all vowel pairs in all prosodic conditions. The effect of boundary strength was significant for all vowel pairs, and significant Post-Hoc Bonferroni results are reported in the last column.	160
6.1	Results on the significance of prosodic boundary strength on F0 measurement points (L1, H1, L2 and H2).	186
C.1	Table showing the mean values of duration and amplitude of closure and release for voiceless and voiced stops.	220
C.2	Results from Binary Logistic Regression analysis testing which predictor values from the duration and amplitude of the closure and release best predict the voicing contrast between voiceless and voiced stops.	221
C.3	Results from Binary Logistic Regression analysis testing which predictor values from the duration and amplitude of the closure and release best predict the voicing contrast between ‘nasal+stop’ sequences perceived to entail either voiceless or voiced stops.	221

List of Figures

1.1	Pre-boundary normalised duration results for two studies investigating the phonetic realisation of break index notations. The left-hand panel shows the duration of pre-boundary rhymes in Price et al (1991), and the one on the right the results from Wightman et al (1992). The vertical axis shows normalised duration, and the horizontal the position within the break index.	7
1.2	Prosodic constituent hierarchies, as proposed by different studies. Figure taken from Shattuck-Hufnagel & Turk (1996, pp.206). Selkirk's model, according to the authors, more closely represents her 1978 and 1981 views.	12
1.3	One possibility of prosodic structure, which has been assumed in a number of phonetic and articulatory studies. Picture taken from Kuzla (2009, pp. 6), and is an adapted version of Beckman and Pierrehumbert's 1986 analysis.	12
3.1	Example of segmentation of items constructed for the investigation of pre-boundary lengthening. Sample taken from a female speaker saying the item [a'ðjakopa ka'lezmata]. Figure shows the names and positions of the syllables, and examples of segmentations. It also points to the position of the boundary. This example is taken from BC4.	68
3.2	Boxplots showing the effect of prosodic boundary strength on duration (ms) of the antepenultimate, penultimate and ultimate syllables of the preboundary word.	73
3.3	Representation of how much the duration of the ultimate syllable was lengthened for each condition	74
3.4	Schematic representation of how each speaker signalled the effect of boundary condition. Domains that stretch within the same line are statistically similar, while domains belonging to different lines are statistically distinct.	77
3.5	Mean duration of pre-boundary rime (i.e. vowel) by speakers MS and CP for renditions BC0-BC5 ($n=144$).	79

4.1	Example of segmentation of items constructed for the investigation of post-nasal stop-voicing. Sample taken from a female speaker saying the item [a'ðapanon ta'belon] (romanisation: adapanwn tampelwn). Figure shows the names and positions of the syllables, and of their segmentation. It also points to the position of the boundary, as well as to positions that were important for segmentation decisions. This example is taken from BC2.	99
4.2	Example of how the relative amplitude of the closure and the burst were measured. This figure shows the window over which the amplitude was measured using PRAAT.	101
4.3	Example of the screen the raters were seeing when running the experiment. Example taken from data for /nk/ classification. Instances classified as either /k/ or /nk/ were later collapsed to a category called 'voiceless', while instances classified as either /ng/ or /g/ were collapsed to a category called 'voiced'. Same categorisation took place for the other two types of stops (/t/ and /p/). The No Answer, Replay, and Oops buttons are seen.	103
4.4	Example of the part extracted from the waveform. The shaded part is the target consonant. In Tier 5 the segmentation from pC to fV shows the part of the sound that was extracted and played to the rater for identification.	104
4.5	Effect of boundary condition on the duration and amplitude of the closure and the release of the phonologically voiceless and voiced renditions.	109
4.6	Percentage of perceptual classification of voiceless instances (all places of articulation and speakers grouped) as voiced or voiceless in each prosodic condition.	113
4.7	Histograms of the relative amplitude of the release with what was classified as voiceless on the left, and voiced on the right. The difference in mean values, and in the amount of voiced instances is apparent. This Histogram plots all measurements from all five boundary conditions.	119
4.8	Hypotheses of how boundary strength will effect the phonetic manifestation of each variable if stop voicing occurs <i>qualitatively</i> (second column), or <i>variably</i> (third column). The first column (separated from the others by a black vertical line) shows the effect found to be exercised by boundary strength on voiceless stops. Based on that expectations are formed. The x axis shows boundary conditions incrementing in size, and the y axis shows Duration (ms for the first two rows) and Amplitude (dB for the last two rows) measurements. Figure is explained further in text.	122

5.1	Example of segmentation of items constructed for the investigation of resolution of vowel hiatus. Sample shows the hiatus in V1V2, and the segmentation of the pre- and post-hiatus segments.	144
5.2	Highlighted in this screenshot is the section extracted from the item [a'ðinato a'poyono] (romanisation: adynato apogono). The vowel pair is [oa] and the preceding consonant and following syllable were extracted and played back to the raters (i.e. they would listen to “toapo”).	146
5.3	Screenshot of the interface for the perception experiment of the [ao] pair.	147
5.4	Percent classification of [ao] and [oa] for each prosodic condition by pooling all raters' results. BC0 is not represented since only two speakers produced instances for that condition.	152
5.5	Mean F1 and F2 values (ERB) at the beginning and end of formant trajectories (and $\pm 1SD$) for each vowel pair.	155
5.6	F2 trajectories for all instances of [ao], [oa], [aa] and [oo] vowel pairs for speaker OM. This graph exemplifies the gradient phonetic output of vowel hiatus resolution in Modern Greek.	156
5.7	Presentation of expectations from boundary strength effect; the first panel shows how formant trajectories of F2 [ao] vowel pairs are expected to look when they meet their typical positions (baseline [a] or [o] positions). The second figure shows a categorical effect whereby one condition is statistically different to the others, within which there is coarticulation allowed, and the third figure shows a gradient effect across domains, whereby more extreme formant values are found higher in the hierarchy, but the effect is not categorical. . . .	158
5.8	Trajectories from beginning to end of the [ao] and [oa] vowel pairs for conditions BC1 to BC4.	160
5.9	Mean duration (sec) for all vowel pairs in all boundary conditions. . .	162
6.1	Tree depicting prosodic structure, as expected by current analyses of the Intonational Approach.	168
6.2	Example of identification of L and H turning points. Tiers 1 and 2 contain information that comes automatically from the MOMEL/INTSINT algorithm. Tier 1 shows the identified turning points, and Tier 2 gives their values from the smoothed out F0 contour. Tier 3 has only the four turning points under investigation, which have been copied by the experimenter. Tier 4 contains the segmentation of the pre- and post-boundary syllables. The depicted item is [a'ðjakopa ka'lezmata], and the example is taken from BC1.	182

6.3	Schematic representation of normalisation algorithm. In the algorithm the raw F0 values are first computed to logarithmic, as shown in the left hand side on the y axis, and then transformed to percent scale.	183
6.4	Mean range values pooled across speakers for each measurement point. Duration is not represented in this graph.	185
6.5	Crude schematic representation of H alignment within the pre-boundary word. All durations are presented proportionally; i.e. the distance from L to H is the 100%, and the duration of the preboundary syllable and alignment of H1 are expressed proportionally to that. For reference reasons the duration of the pre-boundary syllable is also presented (the green line); the alignment of H1 in conditions BC2-BC4 clearly falls within the pre-boundary syllable.	189
B.1	Mean range values for each speaker separately for each measurement point. Duration is not represented in this graph.	219

CHAPTER 1

Introduction

This thesis is concerned with the way spoken utterances are grouped into stretches of speech. The main goal is to investigate the acoustic manifestation of this grouping in order to understand the nature of the structure organising it. This structure is known as the *prosodic structure*, and it is believed that it is made up of a given set of domains, with lower order domains grouping to form higher order ones in a hierarchical manner. Each domain is thought to be qualitatively distinct from all others, and its boundaries to be demarcated in the acoustic signal in a way that distinctively separates it from the other domains of the structure. It is, therefore, believed that the structure is flat, and that there must be some clear acoustic differences across domains, marking the existence of such prosodic phrases.

Recent studies investigating the acoustic and articulatory demarcation of proposed prosodic domains, however, suggest that the anticipated distinctions between purported domains are not always present. On the contrary, they suggest that the acoustic differences found across prosodic constituents often give rise only to variable differences of strength, without support for differences of type. Hence, it can be that the acoustic manifestation of prosodic structure does not reliably support the existence of distinct domains. Such findings cast doubt on the existence of domains that are qualitatively distinct from others in the hierarchy.

This thesis aims at exploring the acoustic manifestation of prosodic boundaries with the goal of testing whether there is evidence for the postulation of qualitatively distinct domains within the prosodic structure on the basis of acoustic data. This

is accomplished by investigating boundary strength effects on acoustic correlates to the structure. In particular, the correlates used are pre- and post-boundary durational marking of constituents, the application of external sandhi, and international marking of constituents. The results will provide insights into whether differences of boundary strength across prosodic contexts give rise to qualitatively distinct boundaries, or only mirror the well-known hierarchy of possible gradations of boundary strength, or a mixture of both effects.

The parallel use of a variety of acoustic correlates to the prosodic structure in one language using instrumental analyses will allow clear insights into whether different domain *types* exist. Additionally, it will show the way prosodic boundaries are manifested acoustically in a variety of boundary strength conditions.

1.1 Prosody and Boundary Strength

It is a general understanding of researchers working on prosody and its interface with other parts of the grammar that units of speech are grouped together to form larger units of speech (e.g. Gee & Grosjean 1983, Halliday 1967, Selkirk 1981, Nespor & Vogel 1986, Pierrehumbert & Beckman 1988). That is, the stream of speech is well-known to be divided into smaller units of some sort. These units are thought to be marked in the acoustic signal in a variety of ways, operating as markers for signalling how phrases are organized. Additionally, it is a well-accepted fact that the disjuncture across some words is in some way *stronger* than across other words, marking the different types of groupings. For example in 1.1 (Cooper & Sorensen 1981, pp. 169) the boundary between the words *John* and *leaves* is intuitively weaker than that between the words *leaves* and *Sue*.

(1.1) If John leaves, Sue will resign.

Strength differences across boundaries have been reported in a number of studies testing the acoustic manifestation of boundaries. A number of acoustic variables have been shown to support the existence of such strength effects, like longer pre-boundary durations, more and longer pauses, higher F0 etc at ‘stronger’ boundaries (Klatt 1975, Grosjean & Collins 1979, Cooper & Paccia-Cooper 1980, Cooper & Sorensen 1981).

It is, therefore, generally acknowledged that spoken utterances are grouped into prosodic phrases of some sort. The boundaries at the edges of those phrases are thought to mirror this grouping in that their acoustic manifestation is in some way *stronger* next to boundaries that are of larger units of speech.

The phonetic implementation of utterances is generally considered as the link providing insights into the mental representation. Therefore, it is generally acknowledged that, in order to understand the organisation of prosodic grouping, it is necessary to use the acoustic signal to deduce the structure. The following sections present proposals accounting for the nature of the prosodic grouping, focusing on the difference between studies suggesting that the sole difference across constituents is that of boundary strength, and studies proposing that the prosodic grouping is regulated by a structure that comprises of a given set of prosodic domain types. We review the acoustic evidence each of those proposals has presented in support of their view, and present how further instrumental research is needed in order to understand the nature of the structure regulating the prosodic organisation of speech.

1.1.1 Direct syntax-prosody mapping

Early analyses on the prosodic grouping of utterances proposed that it is directly regulated by syntax. A *direct syntax mapping* approach was proposed accounting for the ‘stronger’ boundaries discussed in the previous section, and their related lengthening patterns (Klatt 1975, Cooper & Sorensen 1981, Kaisse 1985). This view proposes a direct link between the syntactic structure and the phonological grouping of utterances. Hence, phonological and phonetic phenomena can be explained with reference to syntactic information. It is information from the syntactic constituent structure that is believed to provide the input to the phonological and then phonetic components. For example in (1.1) seen above, it is anticipated that the stronger boundaries are due to some syntactic variable giving rise to the perceived difference in strength, like for example syntactic branching or the type of constituents separating the two words (if they branch directly from an S node, or from VP vs NP nodes etc.).

These strength differences were named *syntactic boundary strength* effects by Cooper & Sorensen (1981) offering one of the very first usages of the term ‘boundary strength’. They proposed an algorithm to predict the strength of the boundaries from syntactic structure. According to this algorithm, there is a set of steps analysing the syntactic constituency of sentences, and giving a value to each word boundary. Importantly, this value is not restricted to a limited number of possible strength gradations, therefore the number of possible boundary strengths is viewed as being of *indeterminate depth*. As Cooper & Sorensen (1981, pp. 182) say “The output of this [final] step is an integer value from 1 to n ”. The direct syntax mapping approach has managed to capture the intuition that there is a range of possible boundary strengths with varying gradations. Given that the depth of the syntactic structure is indeterminate, and that the algorithm for computing syntactic boundary strength is based on syntactic branching, each word boundary can be associated with a variety of possible boundary strength gradations. This view was also supported by Grosjean & Collins (1979), who reported on a correlation between the syntactic break and the magnitude and occurrence of breathing pauses. They also proposed that boundary strength can range in a number of gradations, linking the phenomenon to syntactic influence, and stating that longer and more pauses are to be expected in ‘*more important*’ syntactic breaks. The notion of *longer* durations in *stronger* boundaries again indicates the lack of a given set of possible gradations of boundary strength.

However, research testing the mapping of prosody onto syntax has since shown that the grouping of utterances is not directly regulated by syntax, but there is some other (phonological) structure conditioning prosodic constituency. Research on the grouping patterns of utterances revealed constructions where the syntactic structure could not explain the prosodic grouping, supporting a proposal that there is a different structure organising surface spoken utterances. One of the most well-known examples illustrating a type of syntax-prosody mismatch is shown in example (1.2a and b); (1.2a) shows the syntactic and (1.2b) the prosodic grouping¹ (example

¹More examples and a succinct overview of the issues leading to the postulation of the prosodic structure are presented in Shattuck-Hufnagel & Turk (1996), who discuss cases of discrepancies between the syntactic and the prosodic organization of utterances. Among many examples, they discuss how speakers can have different options for producing prosodically the same syntactic structure (e.g. breaking one utterance to one or more intonational phrases, selecting the tune

taken from Chomsky & Halle 1968, pp. 371).

(1.2)

- a. [This is [the dog that chased [the cat that killed [the rat that ate [the malt that lay in [the house that Jack built]]]]]]]
- b. This is the dog | that chased the cat | that killed the rat | that ate the malt | that lay in the house that Jack built |

Such discrepancies between where the expected prosodic boundary occurs, and where the syntactic boundary occurs, provide clear motivation for the postulation of prosodic structure by highlighting that the structure of spoken utterances is not isomorphic to the syntactic structure (cf. Shattuck-Hufnagel & Turk 1996).

Therefore, research turned to the question of investigating the prosodic structure, that is the structure regulating the prosodic grouping of utterances. While maintaining the belief that the structure is hierarchical, research on prosodic structure aims at identifying the nature of this structure, and its relation to other components of the grammar, like syntax.

The notion of a hierarchically layered prosodic structure that is different to the syntactic one had been explored fairly early on by Halliday (1967), among others. He proposed that the phonological structure of English consists of a set of hierarchically layered domains, which in his analysis are the tonic group, foot, syllable and phonemes, each of which combines to make the next higher level (Halliday 1967, pp. 13ff). This was one of the very first proposals to entail a set of specific prosodic domains making up the phonological structure. In general, prosodic structure has been viewed as the part of grammar that regulates the acoustic and articulatory outputs (Selkirk 1981), with syntax imposing some types of constraints on the possible groupings. It is viewed as:

“[...] a complex grammatical structure that must be parsed in its own right [...] the organizational structure of speech” (Shattuck-Hufnagel & Turk 1996, citing Beckman 1996).

type to be used, or where the focus will be placed etc.), and how some well-formed prosodic structures violate the syntactic structure (like in example 1.2).

This view of prosodic constituency retains the notion of boundary strength, since higher domains are expected to have stronger boundaries than lower domains. The notion of boundary strength is, therefore, now used as a means of inferring prosodic structure and its constituency, not a direct mapping to syntax.

1.1.2 Prosodic structure as consisting solely of differences of boundary strength

Instrumental research on boundary strength and its acoustic manifestation soon revealed that there is a range of possible boundary strengths that can be acoustically and perceptually distinct. As will be presented in this section, acoustic support has been provided to (directly or indirectly) support the view that the prosodic structure might consist (solely) of levels varying in terms of their relative strength.

Two studies in the 1990s, i.e. Price, Ostendorf, Shattuck-Hufnagel & Fong (1991) and Wightman, Shattuck-Hufnagel, Ostendorf & Price (1992), set out to investigate how prosodic constituency is signalled acoustically. Given that they did not want to use any of the proposed theories for prosodic constituency, but wanted the distinctions across domains to be derived from the data, they used a new technique for identifying boundaries. In Price et al. (1991) listeners classified the perceived disjuncture between words in a scale from 0 to 6, with 0 being the weakest and 6 the strongest disjuncture. This was called the *break index* annotation scheme, according to which it was possible to identify the number of distinctions that are perceptually and acoustically distinct (the same technique was used by Wightman et al. 1992). The use of such an annotation scheme shows that both studies did not want to preclude that the prosodic structure is made up of a given set of domains. Rather, this scheme mirrors the belief found in direct syntax mapping approaches, that boundary strength differences could potentially be of indeterminate depth.

Both studies investigated final lengthening, i.e. the elongation of the duration of some linguistic variable (syllable, onset, rime etc) before a prosodic boundary. They both reported longer pre-boundary durations in the vicinity of boundaries classified as stronger on the break index scale. Figure 1.1 shows the results of pre-boundary lengthening (mean and variance) on normalised duration of the pre-boundary rime from Price et al.'s (1991, left panel) and Wightman et al.'s (1992, right panel)

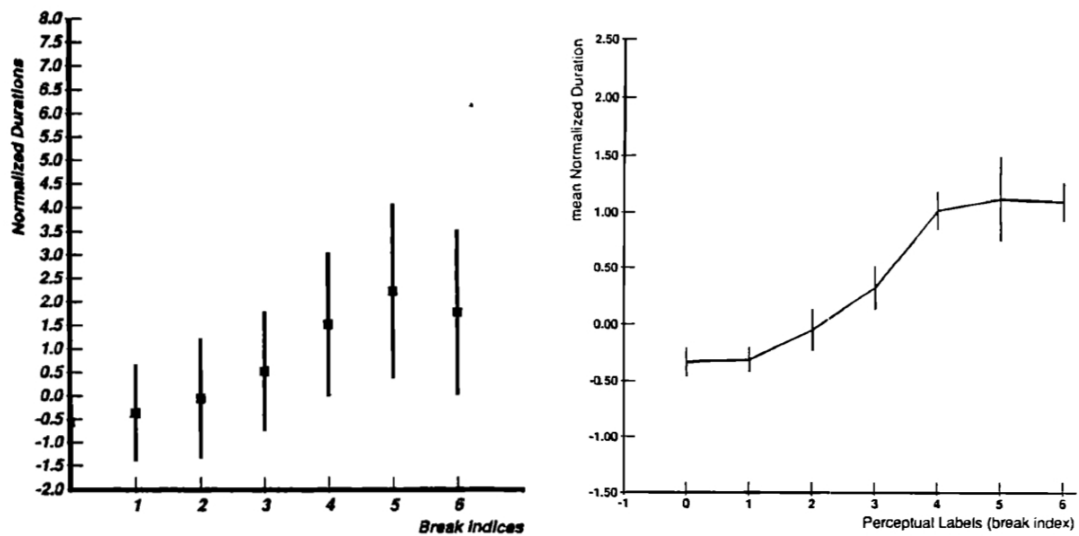


Figure 1.1: Pre-boundary normalised duration results for two studies investigating the phonetic realisation of break index notations. The left-hand panel shows the duration of pre-boundary rhymes in Price et al (1991), and the one on the right the results from Wightman et al (1992). The vertical axis shows normalised duration, and the horizontal the position within the break index.

analysis as a function of the incrementing break index. Price et al. (1991) reported that pre-boundary durations could be grouped in two boundary levels statistically (0-3 and 4-6), and Wightman et al. (1992) reported that solely on the result of pre-boundary durations four statistically distinct levels of boundaries could be discerned (0-1, 2, 3, 4-6). However, both studies show the gradual lengthening in duration, and they clearly depict how listeners can distinguish perceptually a variety of boundary strengths, which, although they overlap statistically, are showing lengthening from one break index level to the other. Moreover, the gradual lengthening effect could be interpreted as support for the view that boundary strength can be of indeterminate depth, although this is not the direct conclusion drawn from the authors.

The potential indeterminacy of boundary depth is also seen in a perception study by de Pijper & Sanderman (1994). They tested Dutch listeners, who were asked to mark boundaries on a 10-point scale, 1 being a very weak boundary and 10 a very strong boundary, in an effort to provide theory-independent evidence for the existence of boundary strength. They showed that there was high inter-rater agreement in their annotations, and that there was a correlation between the number given in

each boundary, and the phonetic cues relating to that boundary. The important point from their study was the high inter-rater agreement amongst naive listeners, even when listening to wavfiles whose speech had been rendered unintelligible. The use of a ten point scale shows that there is a large number of possible gradations that can be identified by listeners of a language.

Additionally, the issue of boundary strength and the question whether it can only signal a given set of domains, or a range of possible gradations was also treated in Ladd & Campbell (1991). They also investigated pre-boundary lengthening and its link to perceived disjuncture, and found that lengthening differences found across conditions can actually give rise to a range of lengthening patterns. Importantly, they directly supported the view that boundary strength distinguishes a large number of gradations, and proposed that those gradations need to be present in our representation of prosodic structure.

The notion of boundary strength was directly used to describe the prosodic structure of utterances in the initial ToBI (Tone and Break Index). In the ToBI annotation system (Silverman, Beckman, Pitrelli, Ostendorf, Wightman, Price, Pierrehumbert & Hirschberg 1992), each utterance is labelled in two separate tiers: the tone and the break index tiers. The break index is a direct use of Price et al.'s (1991) proposal, in which annotators present an analysis of the strength of disjuncture between each orthographic word. This initial type of analysis was meant to mirror the intuition that there are differences of boundary strength that range over a variety of possible values, and which do not necessarily link in a one-to-one relationship with the tone index. This intuition is clearly seen in the continuing use of integer numbers for annotating break indexes, since integers can range indeterminately².

Finally, the necessity for having an indeterminate number of possible boundary strengths that can be acoustically signalled has also been discussed by Ladd (1996) and again in Ladd (2008). On the basis of the sentence in example (1.2), he discussed the possibility of having boundaries between words that are intuitively *stronger* or *weaker* in an 'open-ended scale' (Ladd 2008, pp. 293), supporting the view that

²However, as will be discussed further down, the Break Index analyses of ToBI changed through time to propose that there is a limited set of boundary strength differences, which are inextricably linked to the number of distinct prosodic domains.

there are varying gradations of boundary strength. In particular, he focused on the final part of the sentence ‘This is the dog that chased the cat ...’, i.e. the part *that lay in the house that Jack built*, and reported on his intuition that there has to be another boundary after the word *house*, which however is weaker than all other prosodic boundaries marked in example 1.2.

Other phonetic manifestations of boundary strength come from investigating domain-initial phenomena in recent research, like the articulatory and timing characteristics of sounds at the onset of prosodic domains. Fougeron & Keating (1997) tested the articulatory patterns of /no/ initial and final syllables, and found that the ones following the boundary showed initial strengthening (mainly exhibited in less articulatory tongue contact for post-boundary vowels, with also occasionally stronger articulatory contact for post-boundary nasals), which was cumulative in the vicinity of stronger boundaries. This study was then followed by a number of studies in many languages, showing that, when languages exhibit initial strengthening, it is a cumulative effect. Studies on French, Korean, English, Taiwanese (Onaka 2003, Cho 2005, Hayashi, Hsu & Keating 1999, Keating, Cho, Fougeron & Hsu 2003) among others have since shown that there is such a cumulative effect of initial strengthening, like in pre-boundary lengthening.

All the above discussed studies agree in the fact that there is acoustic evidence for boundary strength differences across prosodic domains. Importantly, there are large differences across the number of classifications of boundary strength conditions used across studies (from 4 or 5 levels up to 10), and in all studies there are differences across conditions suggesting that boundary strength seems to be of variable depth, as was first proposed by theories of direct syntax mapping.

1.1.3 Prosodic structure as consisting of domains of different type

Thus far we have reviewed studies supporting that the prosodic grouping of utterances contains phrases whose boundaries range in their relative strength, and that the available boundary strength differences can assume indeterminate depth. However, researchers working on prosodic structure soon became interested in understanding the exact construction of prosodic groupings. The main questions that arose were:

‘What are the units of prosodic structure, i.e. the linguistically significant sub-parts? What are the principles governing the internal structure and prominence relations within the different prosodic units? How do these prosodic units, or constituents, relate to constituents of syntactic structure?’ (Selkirk 1978, p. 111)

The focus of research, therefore, switched to identifying *specific* constituents making up the structure. The notion of variability of boundary strength that can range in a number of possible levels was given up in an effort to identify a *given set of well-defined domains*, in the sense that each domain can be identified on the basis of some linguistic variable.

This switch in interest made it difficult to reconcile variability in boundary strength, which is what the acoustic manifestation of prosodic boundaries had shown in a variety of previous studies, and the proposal that there is only a limited number of prosodic domains, which also need to be distinctively marked in the acoustic signal. Therefore, the new proposals maintained that each level within the prosodic structure does not only differ in terms of its relative strength to the other domains, but also (and mainly) differs in some important linguistic variable, that provides a qualitative distinction of type across domains.

This view makes an important assumption; that the prosodic structure comprises of a given set of *labelled* domains. Therefore, while like all views on prosodic constituency this proposal accepts the strength hierarchy, it departs from previous analyses in that they attach a specific label and definition to each of the proposed domains. For this reason, these views will be called **labelled views** in the present thesis. The assumption that the prosodic structure consists of a set of labelled domains has been explicitly stated; Selkirk (1981, pp. 381) notes:

‘The phonological units [...] will be referred to as *prosodic categories* [...]. The term “category” is chosen consciously in order to evoke a very apt analogy to the categories of syntax’.

This analysis draws a clear parallelism between syntactic and prosodic domains, suggesting the existence of labelled domains differing in terms of *type*, similarly to syntactic categories.

Proposals entailing a set of labelled domains appeared fairly early on in the investigation of prosodic groupings. For example, Halliday (1967), among others, proposed that there is a given set of domains making up the possible groupings available to English speakers, and that those are hierarchically layered. Recent prevailing proposals put forward within this view are the *indirect syntax mapping approach* (proposing an indirect mapping from the prosodic structure to the syntactic structure for the definition of domains), and the *intonational approach* (using terminology from Jun 1998, referring to research using intonation contours for postulating differences of type across constituents). These proposals make use of phenomena apparent in the acoustic signal to propose the existence of labelled domains, like the application of external sandhi³ and the types of tonal events at the vicinity of prosodic boundaries. Figure 1.2 (taken from Shattuck-Hufnagel & Turk 1996) depicts how Selkirk's (1984) and Nespor & Vogel's (1986) proposals entail a total of five domains above the word level, while Pierrehumbert & Beckman's (1988) comprises of a set of three levels. Similarly, Figure 1.3 (taken from Kuzla 2009) shows a current version of prosodic structure under the intonational approach. It is clear that, while boundary strength differences are still maintained, the important distinguishing factor of these studies is the idea of being able to define a given set of prosodic domains, which are attached to a specific label separating them in a qualitative manner - not only in terms of their relative strength.

Therefore, it is clear how labelled views attach specific labels to each level, and subsequently also link the number of available levels in the structure with the number of possible definitions in terms of syntax-prosody or intonation accounts. The issue of boundary strength assumes a less important role, and the term *labelled views* represents all proposals attaching a specific definition in each prosodic level, suggesting that the differences across constituents are primarily differences of *type*. What those views, therefore, proposed is that there is some intrinsic linguistic variable that creates a qualitative distinction across prosodic domains.

Importantly, labelled views propose that the acoustic manifestation of utterances will provide support for distinctions of *type* across constituents. For example, the

³External sandhi are the phonological processes according to which “segments [...] at the edges of words may undergo changes when words are strung together to produce an utterance” (Arvaniti 2007, pp. 67).

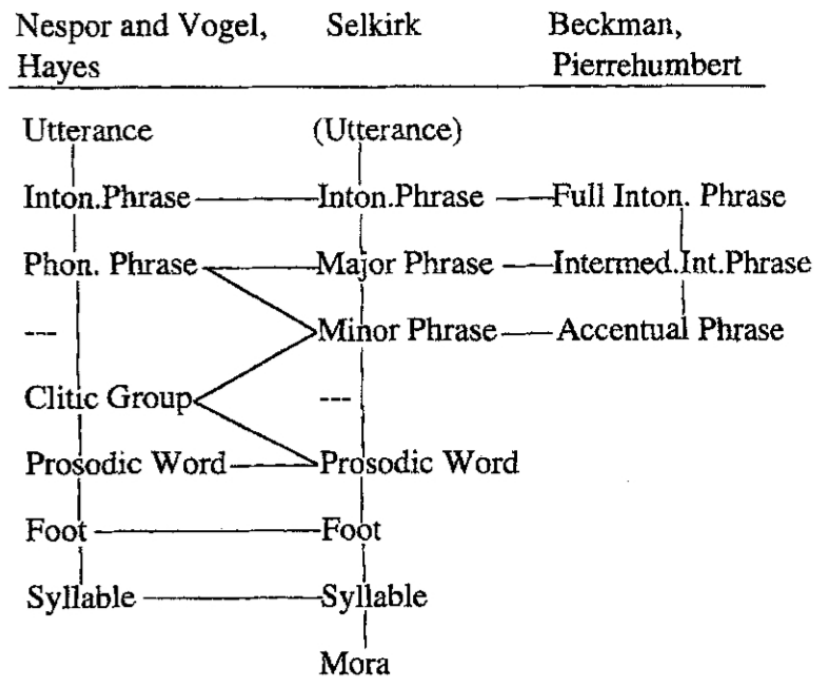


Figure 1.2: Prosodic constituent hierarchies, as proposed by different studies. Figure taken from Shattuck-Hufnagel & Turk (1996, pp.206). Selkirk's model, according to the authors, more closely represents her 1978 and 1981 views.

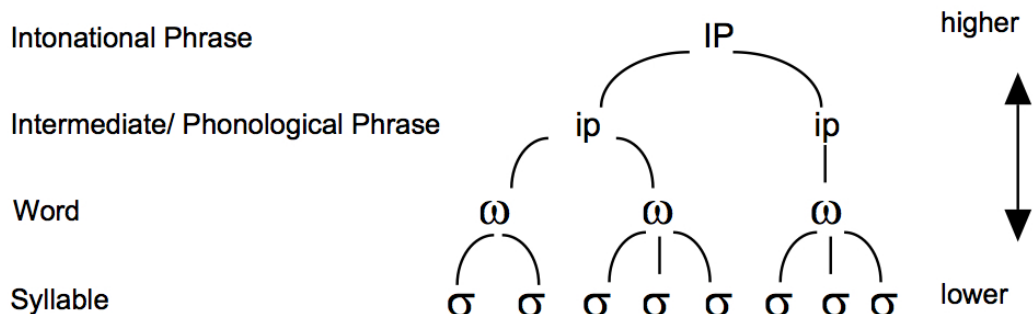


Figure 1.3: One possibility of prosodic structure, which has been assumed in a number of phonetic and articulatory studies. Picture taken from Kuzla (2009, pp. 6), and is an adapted version of Beckman and Pierrehumbert's 1986 analysis.

indirect syntax mapping approaches motivate prosodic domains above the word level mainly on the basis of the application of external sandhi. That is, having a sandhi process being blocked across domains provides evidence for the existence of a prosodic constituent. The intonational approach mainly motivates distinct domains on the basis of intonation contours, whereby different intonational events

are linked to different levels in the structure. The application of sandhi phenomena and intonation contours assume a crucial role in the definition and/or identification of prosodic constituents. Acoustic differences across prosodic levels are taken to offer support for the postulation of prosodic domains, and are all expected to act as markers of prosodic constituency.

While analyses of sandhi and intonation contours within labelled views assume that there is solid phonetic evidence in the acoustic signal for postulating distinct prosodic domains, more recent studies question the existence of acoustic and/or articulatory support for differences of type, and warrant a re-investigation of the assumption that the prosodic structure is made up of labelled domains. The following section reviews evidence from recent acoustic and articulatory studies motivating the question whether there is indeed support on the acoustic manifestation of prosodic boundaries for postulating different *types* of domains.

1.2 Do labelled domains exist?

1.2.1 Evidence against labelled domains: from sandhi

In the previous section it was hinted that the application of external sandhi has assumed an important role in identifying domains in the prosodic structure. It is assumed that these phenomena will provide support for the existence of some domains in the structure. The main focus of this section is to identify the inadequacy of traditional impressionistic analyses of external sandhi to provide solid support for the existence of labelled domains. It is shown that instrumental analyses of sandhi phenomena suggest that very often only differences of strength are mirrored in the application of sandhi, quite similarly to the application of pre-boundary lengthening. Such a finding would question proposed differences of type across domains. The research presented in this and the following sections also provides the motivation behind the present thesis.

External sandhi has been crucial in the postulation of different types of prosodic domains, being one of the main segmental processes providing support for the existence of labelled domains. However, recent studies raise questions regarding their applicability as correlates to labelled domains. Two issues will be discussed in this

section: (i) the fact that analyses of sandhi phenomena within labelled views have used *impressionistic data*, often classifying segmental processes as being sandhi (therefore being the result of a phonological alternation), while subsequent instrumental research has shown them to be the result of coarticulation (therefore with the possibility of having a variety of different productions), and (ii) the fact that the effect of boundary strength on sandhi has been assumed to be categorical distinguishing qualitatively different domain types, when current research has shown that it can be variable across prosodic domains⁴. Both topics make it imperative that sandhi phenomena are reviewed as potential markers of prosodic constituency.

Regarding the first point, external sandhi has mostly been treated using impressionistic analyses, which rely on the experimenter's perception and classification of the resulting phonetic output. While using impressionistic analyses can be fruitful for some experiments, it is possible that perceptual classifications of whether a phenomenon takes place or not can be biased, in that the continuous acoustic signal is interpreted in a categorical manner. Gibbon & Scobbie (1997) compared impressionistic transcriptions and articulatory analyses using EPG of speech from children with phonological disorder showing that the two types of analyses can yield different results. Their research provided concrete evidence on how on the basis of articulatory analyses of clinical patients' speech they found blending of production, when analyses using impressionistic transcriptions gave rise to categorizations of output.

Similarly in the case of sandhi, a phenomenon that might not be categorical could be transcribed impressionistically as categorical due to the mediation of the experimenter. This is important for the present discussion since labelled views have assumed several external sandhi phenomena to be phonological, and therefore to either apply or not, which in subsequent research have been found to be coarticulatory. For example, vowel hiatus resolution in Modern Greek has been traditionally transcribed as a sandhi phonological rule (Kaisse 1985, Nespor & Vogel 1986, Condoravdi 1990, Drachman & Malikouti-Drachman 1999), and to mainly be resolved by deletion of the first vowel of the vowel pair (we will elaborate on this

⁴More on the definition of the nature of the processes themselves, as well as of the possible boundary strength effects on them in Sections 1.2.3, 1.2.3.1, and 1.2.3.2.

further in Chapter 5). Recent research testing the phonetic characteristics of vowel hiatus resolution in Modern Greek, however, has shown that it involves varying degrees of coarticulation (Pelekanou & Arvaniti 2001, Baltazani 2006, Kainada 2007).

There is also much research on other languages and other processes that has brought to question the issue whether traditionally considered sandhi phenomena are rightly viewed as phonological processes entailing categorical segmental alternations. A series of studies in the 1990s using acoustic and articulatory data showed how processes impressionistically transcribed as categorical have gradient phonetic outputs. Nolan, Holst & Kühnert (1996) investigated the articulatory characteristics of [s] to [ʃ] place assimilation in English in a variety of contexts (like for example between the words *miss you*) using electropalatography (EPG), and reported that the articulatory patterns found ranged in a continuum of productions with intermediate gestural results being the most common, pointing to a gradient phenomenon rather than a categorical phonological process. Similar findings were reported in Zsiga (1997), who investigated vowel hiatus resolution in Igbo vowel pairs, and reported that the phenomenon is not phonological, contrary to traditional analyses. On the other hand, Ladd & Scobbie (2003) showed instances of consonant degemination and lenition in Sardinian Italian for some speakers, which appeared to be resolved categorically, complicating the question whether sandhi phenomena are the results of phonological cognitive planning or of coarticulation.

Moreover, disagreement can be seen on whether sandhi processes are phonological even between studies using acoustic data. For example, /s/-voicing in Modern Greek, a process impressionistically described by Nespor & Vogel (1986) as a categorical process changing /s/ to [z] before voiced consonants, has also been tested by more recent studies using acoustic data. Pelekanou & Arvaniti (2001) suggested that it has a phonetically gradient output, while Tserdanelis (2004) reported that it has an acoustically categorical output. Such discrepancies can arise from using very different methodologies, or a limited set of prosodic conditions, speakers, items, or repetitions. Baltazani (2007b) identified, for example, this mismatch between studies, and by using a variety of prosodic conditions and speakers she proved that the particular process has a phonetically gradient output. However, such mismatches even between acoustic studies highlight the need for careful investigation of the

nature of those processes, so that conclusions regarding whether they can motivate prosodic domains can be reliable⁵.

Therefore, there is a distinction between sandhi processes and coarticulatory processes, which is important with respect to providing motivation for prosodic domains. The reason it is important is that there are different expectations formed for sandhi vs coarticulatory processes with respect to the types of boundary effects a researcher will find. If it is a sandhi process, then prosodic domain identification takes place by investigating whether the application of the phenomenon is allowed or not across prosodic boundaries, or how many times it takes place across each type of boundary (if the effect is found to be variable). On the other hand, if a process is coarticulatory, the effect of the prosodic boundary can no longer be identified on the basis of whether the process takes place, since it is not a categorical phonological process; rather, it is measured on phonetic variables expressing the amount of coarticulation. For example, in vowel hiatus resolution the formant trajectories from the first to the second vowel of the vowel pair will be an indicator of how much coarticulation has taken place. In order for a coarticulatory phenomenon to act as marker of prosodic constituency it means that coarticulation will be allowed across some prosodic boundaries and completely blocked across others, while if the effect is variable across domains the degree of coarticulation will vary across conditions. Therefore, in the case of a coarticulatory phenomenon what is measured is some acoustic variable indicating how much coarticulation is taking place.

To conclude with this first issue, the postulation of labelled domains has relied heavily on impressionistic analyses of sandhi phenomena. Current findings from instrumental studies suggest that some of those phenomena are in fact the result of varying degrees of coarticulation. Hence, conclusions regarding prosodic constituency based on their application need to be re-addressed.

With respect to the second point mentioned earlier, one of the implicit assumptions of impressionistic analyses of sandhi across prosodic conditions has been that boundary strength will affect the application of sandhi processes across domains

⁵Recent studies have also been performed in coarticulatory processes in Cypriot Greek. For example, vowel lengthening (Eftychiou 2007) and /s/-voicing (Eftychiou 2008) have both been shown to be gradient processes.

qualitatively. This means that it is assumed that the phenomenon will either apply or be blocked in a categorical fashion across domains giving rise to qualitatively distinct domains. However, it is equally possible that the effect of boundary strength will be variable, therefore the question arises whether proposals for labelled domains are biased by the assumption of categoriality, or whether there are indeed processes affected by boundary strength in a qualitative manner.

This question again stems from recent findings in the literature reporting that sandhi processes have been found to undergo a variable boundary strength effect; for example, Baltazani (2007*b*) on /s/-voicing reported that the distribution of instances where /s/-voicing took place was variable across prosodic domains, with fewer instances higher in the structure.

Importantly, since some segmental alternations have been found to be coarticulatory, boundary strength effects on the amount of coarticulation could also act as a correlate to prosodic constituency. For example, vowel hiatus resolution in Modern Greek has been found to be the result of coarticulation, but the effect of boundary strength on it has been mixed across studies. Pelekanou & Arvaniti (2001) report that coarticulation in vowel hiatus resolution is completely blocked across intermediate phrases, while Baltazani (2006) reports that it is blocked across intonational phrases. These discrepancies could arise from different methodologies (for example Baltazani 2006 included items with pauses separating the two vowels, which, if there were more of those in the intonational level could create an artifact in the finding that intonational phrases block coarticulation), as well as from differences across domains and their construction.

To sum up, instrumental studies on external sandhi have provided two reasons for questioning whether labelled views have correctly proposed the existence of labelled domains within the prosodic structure. First, testing whether sandhi phenomena are indeed phonological or the result of coarticulation, and second, whether the boundary strength effect on their application is categorical or variable across conditions. It is the matter of instrumental analyses to test whether there are segmental phenomena that can provide support for labelled domains.

One interesting observation from current research on sandhi phenomena on Modern Greek with respect to the effect boundary strength exercises on them is that there hasn't been any phonological sandhi process reported that is proposed to be influenced variably by boundary strength. All variable boundary strength effects discussed in the literature are observed on coarticulatory processes, like /s/-voicing. The distinction between the nature of the sandhi process and the effect boundary strength exercises on it is important to maintain, but has not been clear in much research. For example, Baltazani (2007*b*) specifically mentions: "One important implication of these results [the finding that a process is not phonological] is that sandhi is not a reliable criterion for detecting prosodic boundaries since its outcome is gradient". This clearly suggests that, if a process is found to be coarticulatory, it cannot offer motivation for postulating specific labelled domains in the structure. However, we believe that this is not the case, and that coarticulatory processes can also be influenced qualitatively by boundary strength by having all coarticulation blocked across some domains and not across others. While it is appealing to suggest that there is a link between a coarticulatory process and a variable prosodic effect, in the sense that, if a process is coarticulatory then the boundary strength effect will necessarily be variable, this should be the result of acoustic and articulatory investigation. It is one of the goals of this thesis to separate the nature of the phenomenon from the type of effect exercised on it by boundary strength (see Section 1.2.3 on how the two issues are kept distinct, and for terminology used in this thesis).

Throughout this section most examples presented have been derived mainly from Modern Greek. This was deliberate, in order to highlight the importance Modern Greek sandhi has played in the shaping of theories of indirect syntax mapping. Modern Greek is a language rich in sandhi processes (see e.g. Tserdanelis 2004, Arvaniti & Baltazani 2005, for lists of attested sandhi phenomena in Modern Greek, which do not claim to be exhaustive), the analysis of which has yielded much disagreement among researchers. The picture regarding the phonetic and phonological nature of sandhi processes in Modern Greek and the effect of boundary strength on them has been shown to be unclear. The plethora of sandhi processes available,

and the discrepancies reported across studies, make this language very appealing for investigating the question at hand.

The main issue discussed in this section, therefore, has been the inadequacy of traditional impressionistic analyses of external sandhi to provide solid support for the existence of labelled views. The following section focuses on the use of intonational patterns for the postulation of labelled domains, and how several assumptions within that research also warrant re-investigation.

1.2.2 Evidence against labelled domains: from intonation

The focus of the present section switches to the use of intonation for proposing a set of domain types within the prosodic structure. The discussion will revolve mainly around studies using intonation for postulating prosodic domains, e.g. Pierrehumbert and Beckman's work and subsequent research within the ToBI system. That is, we do not refer to indirect syntax approaches which, although they acknowledge the delimitative function of intonation, they only consider it as relevant for some of their proposed domains.

It was shown in Figure 1.3 that the most commonly accepted version of prosodic structure within studies using intonation contours for the identification of domain types entails three levels above the word: the prosodic word, the intermediate phrase and the intonational phrase. The prosodic word is qualitatively different to the intonational and intermediate phrase because it is identified on the basis of having a primary stress, and is not considered to form a prosodic phrase on its own (although it can). It is also often associated with a pitch accent. The difference between the intermediate and intonational phrases is more difficult to address; the assumption is that the factor differentiating those two levels is the type of tone occurring at the end of each domain; for example in their analysis in Modern Greek Arvaniti & Baltazani (2005) mention that:

‘For the intonational analysis of Modern Greek we recognize three types of tonal events: *pitch accents*, which associate with stressed syllables, and two types of phrasal tones, *phrase accents* and *boundary tones*,

which associate with the boundaries of intermediate and intonational phrases respectively.’ (Arvaniti & Baltazani 2005)

It is, therefore, accepted that each phrasal level is associated with a particular tonal event, and assuming that each tonal event is qualitatively different in some way, the domain it associates with is also assumed to be of different type.

In this section we will discuss how the above description needs revisiting based on two issues: (i) there is an expedient assumption in this line of research that differences of boundary strength are secondary to differences of type, suggesting that acoustic differences of boundary strength are not to be represented in the structure, and (ii) the prosodic domain types proposed for the phrasal level (i.e. the intermediate and intonational phrase) do not have clear definitions, and their acoustic differences resemble more differences of strength within the same constituent, rather than differences of type.

Starting with the first issue, the motivation for postulating prosodic domains is the associated tones delimiting their edges. Initial analyses of tones and prosodic levels (like Price et al. 1991, who proposed the break index analysis, and proto-ToBI proposals) kept the identification of tones separate to the identification of levels. That is, the former was meant to describe the intonational contour, and the latter to provide insights into the prosodic grouping. Hence, initially it was not proposed that tones and prosodic domains were in a one-to-one relationship, with the number of existing tones predicting the number of existing prosodic domains. On the contrary, the break index entailed seven levels, capturing a wide range of possible degrees of separation across domains.

Further research and the proposal of the ToBI system, however, subordinated the break index analysis to the tonal analysis by attaching the tones to specific levels. The number of possible break indices was lowered, mirroring the number of available tones. Therefore, the numerical break index was impoverished, and a link between break index levels, specific tones, and labelled domains was postulated (confer for example Price et al. 1991, Wightman et al. 1992, who used seven levels of break index, and Ladd & Campbell 1991 who showed how more levels are marked durationally than there are domains). For example, a break index of 1 is expected to

occur across two word boundaries in fluent speech, break index level 3 marks the intermediate phrase which co-occurs with a phrase accent, and a break index of 4 co-occurs with a boundary tone and with the marking of an intonational phrase. While the postulation of different tonal events might be viewed as cue to labelled domains, there is no reason to propose that the number of possible available domains mirrored in the break index analysis should match that of the available tonal events. Beckman & Elam (1997) clearly show how the BI analysis was subordinated to the To analysis:

‘These two break index strengths [i.e. 3 and 4] are equated with the intonational categories of intermediate (intonation) phrase and (full) intonation phrase. Thus, whenever the tonal analysis indicates a L- or H- phrase accent, the transcriber should decide where the end of the intermediate phrase marked by this tone label is and place a 3 on the break index tier to align with the orthographic label for the last word in the intermediate phrase. Similarly, whenever the tonal analysis indicates a L% or H% boundary tone, the transcriber should place a 4 on the break index tier at the end of the last word in the intonation phrase.’ (Beckman & Elam 1997, pp. 33)

It is, therefore, clear that the subordination of the break index to the tone analysis potentially takes away some of the phrasing information that was available from a richer inventory of break indices. Importantly, it imposed the notion of labelled domains on the structure, which was not part of the initial prosodic phrasing analyses. However, that was an expedient decision, which might not provide the optimal way of capturing prosodic constituency. It is a decision that limits the number of prosodic domains available, and importantly it precludes that differences across domains are of quantitative nature, an assumption that was not necessarily present in the break index view of boundary strength.

Moving on to the second issue: linking tones and break index to specific domains meant that specific domains are expected now to differ in terms of some linguistic variable, not only in terms of strength. This means that, on the basis of suprasegmental patterns, there needs to be some sort of acoustic evidence supporting the

existence of qualitatively distinct domains. However, both the *definition* of the difference between phrase accents and boundary tones, and their actual *acoustic manifestation* cast doubt on the claim that the two levels are distinct in terms of type.

Firstly, regarding the definition of the two types of tones, the difference between phrase accents and boundary tones has been said to lie in variables like F0 scaling, final lengthening, and pausing, and in particular rely on a notion of *comparison* between the two; that is, it is expected that the intonational phrase will show *more* final lengthening, *longer* pauses, *higher* scaling etc (Selkirk 1978, Nespor & Vogel 1986, Pierrehumbert & Beckman 1988, Arvaniti & Baltazani 2005). Therefore, the difference between the two types of events entails the notion of *variability across domains*, which, however, is primarily (but not necessarily so) a cue to differences of strength across domains, rather than type. Therefore, the definition of the distinction between the two types of events itself makes it hard to identify the differences across events and domains.

Moreover, the actual acoustic manifestation of the two purported domains has been examined by recent research and has shown that indeed the distinctions are not categorical as would be expected in a view postulating differences of type. Chavarria, Yoon & Cole (2004) and Yoon, Kim & Chavarria (2004) set out to investigate acoustic differences across tones marked as phrase accents vs boundary tones; they focused on differences between L- phrase accents and L-L% boundary tones from the English Switchboard and the Boston University Radio News corpora. They found significant differences in the duration of the pre-boundary rhyme, in the sense that tones identified as boundary tones were associated with longer rhymes than the rhymes that were associated with phrase accents. However, they found no differences in F0 between tones they had marked as L- phrase accents vs L-L% boundary tones. They concluded that there are two distinct prosodic level types, but their conclusion stemmed only from the differences found in final lengthening, and was not supported by differences between the actual phrase accents vs boundary tones. This by itself poses serious questions for the postulation of two distinct tonal events, and questions the view that the two levels are qualitatively distinct. While a finding of variable boundary strength effects does not exclude that there are underlyingly

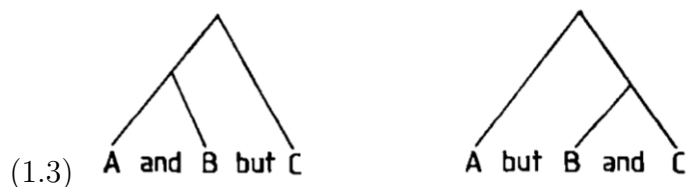
differences of type in the prosodic structure, it also cannot provide solid support for postulating such distinctions.

This study highlights the fact that researchers are inclined to regard any statistical differences as mirroring differences of type across prosodic domains. This assumption can lead to mistaken postulation of distinct prosodic domains that are actually not present in the prosodic structure, but are varying solely in terms of relative strength (see Section 1.2.3 for a discussion on whether statistical differences can be treated as cues to differences of type). The findings from Chavaria et al. seem to support a view of variable differences of boundary strength across the investigated conditions, which is the exact opposite conclusion to the one they reach. Moreover, even their durational data, upon which they based their conclusion, entail a significant amount of overlapping across purported levels.

The general subject matter of this section has been the question of whether internationally and durationally there is support for a given set of prosodic domains that are marked via a qualitative boundary strength effect. Ladd (1988) investigated a related question, that provides further insights to the nature of prosodic constituency on the basis of its acoustic signalling. Ladd's investigation questioned a prevalent assumption within the prosodic structure, that of the Strict Layering Hypothesis, which states that each prosodic domain needs to be parsed to the immediately lower level with no recursion allowed. In this view, differences of boundary strength within the same type of domains are not allowed, and the only types of effects anticipated are those giving rise to differences of type. However, Ladd's (1986, 1988) papers presented theoretical and experimental evidence supporting the view that the acoustic manifestation of boundary strength captures more distinctions than the available labelled domains, and that those distinctions appear to be differences of strength across constituents of the same type.

Ladd (1988) provided phonetic evidence from pitch accent scaling as a cue to differences of prosodic boundary *strength* in an experiment that tested sequences like in (1.3). The experimental question tested whether differences in the intonation and timing patterns across the two types of sentences are to be viewed as differences of *strength* involving the same prosodic boundary which is signalled with *more* or

less of some notion of an acoustic cue, or whether they represent differences of *type*, giving rise to the postulation of two different qualitatively distinct prosodic domains:



[Allen has more popular policies, and Warren is a stronger campaigner], [but Ryan has a lot more money].

[Allen has more popular policies], [but Warren is a stronger campaigner, and Ryan has a lot more money].

He investigated differences in the topline intonation patterns (F0 peaks) of the *and/but* and *but/and* structures and found that each clause was marked as an intonational phrase on its own using declination at the beginning of each clause, but by looking at the topline patterns of the intonation contours across the two sets of sentences he found that the F0 was higher and pause duration was longer before *but* boundaries, suggesting that the boundary before *but* clauses was *stronger*. He proposed as a solution that recursion should be allowed within the prosodic structure, which would represent that the different types of clauses are of the same type, but the break before *but* clauses is ‘stronger’ (calling the resulting structure a Compound Prosodic Domain).

Similar findings were reported in Ladd & Campbell (1991), who investigated pre-boundary durations and reported that the number of statistically distinct domain types exceeded that of intonationally defined domains, another finding in support of his analysis involving recursion.

The finding that differences across such types of syntactic constructions might not represent categorical differences of domain *type*, but rather differences of boundary strength has been replicated by some studies since, and it has been suggested that gradient variability might stem from relative scaling of boundary strengths to each other (Féry & Truckenbrodt 2005, Abney 1992, Wagner & Gibson 2006). While such

studies take us a long way into understanding that some type of recursion might need to take place within the structure, it is vital that the acoustic manifestation of such effects is investigated in a variety of prosodic contexts, and not only in types of sentences like those in Ladd (1988).

Ladd's analysis is very important, since it is one of the very few studies suggesting that differences of strength need not be linked to differences of type in a one-to-one relation, as supported by the labelled view. Rather, Ladd's view departs from this traditional analysis in the important way that, while labelled views are still allowed (and needed) in the structure, each level is allowed to be broken down to levels of the same *type*, but of different *strength*, capturing the intuition present in the break index that there are perceptually more distinctions across prosodic levels than there are captured in the intonational approach. Moreover, by keeping differences of strength and type as possible options, the two are viewed as two different properties of the structure, which can co-occur.

Therefore, Ladd's proposal combines differences of type and of strength in the representation of prosodic structure, providing another possibility for research investigating how acoustic manifestations of boundary strength mirror the nature of prosodic constituency. In the present thesis it is not assumed that Ladd's Compound Prosodic Domains proposal is the only solution for combining findings of qualitative differences and variability across boundary conditions. In Compound Prosodic Domains recursion only takes place once, with the domain repeating itself at the lower level; however, it might be the case that there is only a very limited number of distinct labelled prosodic domains, and that the recursive ability of those could mirror more strength gradations. The subject matter of the present thesis is not to test this hypothesis, but to investigate whether there is reason to postulate labelled domains in the structure. However, Ladd's proposal offers a possible solution to parallel findings of qualitative and variable boundary strength effects, and provides one of the possible explanations for mixed results, and findings from this thesis will be discussed with respect to his proposal in Chapter 7.

To summarize, this section added to the discussion of how proposals using tonal events for the postulation of labelled domains need to be tested acoustically. This

discussion adds to the one from previous sections on how sandhi phenomena have also been shown to need revisiting before being used as proof for the existence of labelled domains. Throughout these two sections it has been shown how the issue of boundary strength effects on correlates to prosodic constituency is vital for understanding the nature of the structure regulating prosodic grouping. In the following section we review the types of findings from boundary strength effects on acoustic correlates to prosodic structure that can be reliably taken to provide support for qualitative differences of *type*. Alongside this, we review the terminology used in the rest of this thesis.

1.2.3 Terminology and acoustic evidence for domain types

In the course of this introduction it has been shown that there is a number of correlates to the prosodic structure, e.g. external sandhi, occurrence and length of pauses, pre-boundary lengthening, F0 alignment and scaling etc. This section focuses on one important issue: the type of acoustic evidence anticipated when proposing the existence of different domain *types*.

Past research on postulating specific labelled domains within prosodic structure has often used confusing terminology for describing boundary strength effects, or has used similar terminology for very different matters. For example, a segmental alternation can be said to be a sandhi phenomenon, or the result of coarticulation. Much research has used the terminology of *categorical* vs *gradient* meaning that when the segmental alternation is complete from one sound to the other it is categorical, while if the alternation is not complete but assumes intermediate values from one sound to the other, it is gradient. However, the same terminology has also been used to describe the effect of boundary strength on such processes; that is, a boundary strength effect can be called again either *categorical* or *gradient*. A categorical boundary strength effect is one that blocks an effect from happening across one domain, and allows it across others, while a gradient effect gives rise to gradually less instances of a phenomenon taking place across prosodic domains. The use of the same terminology for the two completely different issues has led to confusion, with researchers often making the assumption that only what is called a categorical phenomenon can undergo a categorical boundary strength effect, while

a gradient phenomenon (like coarticulation) can't. In this thesis we will keep the two issues distinct by using different terminology for referring to each phenomenon; the acoustic output of segmental alternations will be referred to as taking place either categorically or gradiently (without necessarily providing support for whether the process itself is the result of a phonological alternation, or the result of coarticulation), while the effect of boundary strength on the different correlates will be referred to as qualitative or variable. The next two sections elaborate more on these distinctions.

The second, and main, issue is to clearly state the types of findings anticipated if we are to propose that the acoustic signal provides support for the existence of different domain *types* within the structure. It is often not clear in past research what kinds of effects are thought to reliably mirror differences of type across purported constituents, and it is often the case that statistically significant differences across boundary conditions are taken as a marker of differences of type across constituents (section 1.2.3.3 specifically discusses why statistical significances can't always be used to postulate distinct domain types within the structure).

1.2.3.1 Nature of process investigated

The correlates of prosodic structure used in this thesis are external sandhi, pre-boundary lengthening and F0 alignment and scaling. Pre-boundary lengthening and F0 alignment and scaling are well-known to be *gradient*, in the sense that their phonetic output ranges in a continuum of possible values. Therefore, both of those processes are regarded as gradient in this thesis.

External sandhi, on the other hand, is a more complex category. Sandhi can be said to be categorical, if the result of the segmental alternation is complete, changing one sound completely, for example if vowel deletion takes place it is expected that in a vowel hiatus pair [ao], if V2 gets deleted the resulting sound will be [a] and its durational and formant characteristics will resemble those of an [a] sound. However, a segmental alternation can also be gradient, varying in a continuum of possible outcomes. These two possibilities mirror the difference between whether a phenomenon traditionally viewed as a sandhi process is actually phonological, or whether it is the result of coarticulation. The specific question of what is viewed as

a phonological process, and what as the result of coarticulation, and whether the two are distinct, is not the matter of the present thesis. We, therefore, take the very general stand that sandhi is a phonological phenomenon whose acoustic output can fall into two neatly distinct categories, i.e. the category with the sandhi phenomenon taking place and the one where it doesn't. Coarticulation, on the other hand, is to be identified on the basis of variable phonetic outputs across items. We are fully aware that variable phonetic outputs do not exclude the possibility of having two distinct categories underlyingly, but this issue is not the matter of our investigation.

Therefore, we will be using two terms: a *sandhi* process is taken to be phonological. A *coarticulatory* phenomenon is taken to be a segmental alternation that can vary in a continuum of values, and can range across renditions within speaker. The acoustic output of such processes in this thesis is taken to be either categorical or gradient. In the first instance, it is expected that the resulting acoustic outputs will be able to fit into two distinct categories, i.e. the one where the sandhi phenomenon takes place and the one where it doesn't. In the second case, the acoustic output can range in a variety of values across renditions. In this thesis we will mainly investigate whether the acoustic output of several segmental processes is categorical or gradient, and compare this finding to past research. It is not the goal of this thesis to provide an answer as to whether phenomena showcasing categorical effects are to be viewed as phonological (i.e. sandhi), while phenomena that have gradient outputs are to be viewed as phonetic (coarticulatory). While we will discuss the findings under this scope, the main goal is to identify whether the acoustic output is categorical or gradient, in order to be able to test the effect boundary strength exercises on it too. For purposes of convenience, a categorical finding is generally viewed as a sandhi phenomenon in the traditional view (e.g. in the sense that Nespors and Vogel (1986) would have identified it too), while a gradient finding will be generally viewed as the result of coarticulation.

The reason why it is important to first identify whether a process is a sandhi-type phenomenon⁶ or the result of coarticulation is that the identification of boundary

⁶The term 'sandhi-type phenomenon' will be used for phenomena which seem to behave like sandhi processes, but for the nature of which acoustic or articulatory research has not concluded.

strength effects will differ across sandhi vs coarticulatory phenomena. If a phenomenon is found to be a sandhi process, then the effect of boundary strength on it can be measured by investigating how many times it takes place across boundary conditions. If the phenomenon is said to be coarticulatory, the effect of boundary strength then has to be measured on the amount of coarticulation taking place across boundary conditions. There is, hence, a difference in what is being investigated depending on the nature of the process; in the case of sandhi we investigate the distribution of how many times the phenomenon occurs across conditions, while in the case of coarticulation, we investigate the acoustic (or articulatory) behaviour of some variable across conditions, testing the amount of coarticulation.

1.2.3.2 Types of boundary strength effects

In identifying whether prosodic structure consists of distinct domain types, it is important to define what types of boundary strength effects are taken to mirror differences of *type* across boundaries. In this thesis we divide possible boundary strength effects to two categories: the **qualitative** boundary strength effect, and the **variable** boundary strength effect⁷.

A **qualitative** boundary strength is when a correlate is in some linguistically relevant way different across boundary conditions. This effect is taken to mirror a difference of type across boundary conditions, hence to provide support for the postulation of qualitatively different types of prosodic domains. In the case of sandhi phenomena, a qualitative boundary strength effect is easily identified: if a sandhi phenomenon is blocked across some conditions and allowed across others, then we can identify a qualitative boundary strength effect. If the phenomenon is blocked across one domain, it is assumed that it will also be blocked across all higher domains too, giving rise to a binary distribution.

This term will be used throughout the thesis, contrasting it to the term ‘sandhi’, which is expected to be a phonological alternation taking place categorically.

⁷The use of the particular terminology (qualitative vs variable) is contrasted to the use of the terms categorical vs gradient, which in this thesis will be solely used in the context of the acoustic output of segmental alternations. This, as said before, is done in order to maintain the distinction between the acoustic output of segmental processes and the effects of boundary strength on those processes as distinct as possible.

In the case of coarticulatory phenomena, a qualitative boundary strength is again easily identified on the coarticulatory patterns across boundary conditions. A qualitative boundary strength effect is one that categorically blocks coarticulation across a boundary, and allows it across others. For example, a qualitative boundary strength effect on vowel hiatus resolution is one that allows coarticulation in some domains giving rise to variable formant values at the beginning and end of the formant trajectories, but blocks coarticulation across some domains, therefore the formant trajectories at the beginning and end of the vowel pair resemble the typical values for the vowels.

A qualitative effect on sandhi and coarticulation is easily identifiable. Pre-boundary lengthening and F0 alignment and scaling, however, are harder to address, since they are inherently gradient variables. The most reliable cue for a qualitative difference across domain types can be found when lengthening or scaling occur on different linguistic variables across prosodic conditions. For example, hypothetically pre-boundary lengthening could occur on the pre-boundary rhyme for some levels, and on the final stressed pre-boundary rhyme for others. This would give rise to a qualitative distinction across prosodic conditions. F0 alignment in Modern Greek has been shown to provide such a qualitative distinction between the word and the phrase levels; the alignment of the pre-boundary H tone is expected to associate with the final stressed syllable in the prosodic word level, since it is linked to a pitch accent, while in higher domains it is expected to associate with the prosodic boundary, since it is linked to edge tones. Therefore, such inherently gradient variables can also provide the basis for postulating distinctions of type across prosodic levels.

The other type of possible boundary strength effect is the one named in this thesis a **variable** boundary strength effect. A variable effect gives rise to gradually *more* or *less* of some phenomenon higher in the structure. Generally, it is expected that there will be *more* pre-boundary lengthening higher in the structure, and maybe *higher* F0 scaling. With respect to sandhi phenomena, it is expected that there will be gradually *less* instances of a sandhi phenomenon occurring higher in the structure. For example, if stop-voicing is a sandhi phenomenon in Modern Greek and it is influenced variably by boundary strength, then we expect to see more

instances of stop voicing lower in the structure. With respect to coarticulation, a variable boundary strength effect is one that allows *more* coarticulation lower in the structure. Taking again vowel hiatus resolution for an example, if it undergoes a variable boundary strength effect, formant trajectories are expected to be more typical-like higher in the structure, since less coarticulation will be allowed there, and more ‘blended’ lower in the structure, where more coarticulation is allowed.

Past research on boundary strength effects on sandhi phenomena often did not acknowledge the possibility that boundary strength can affect the distribution of a sandhi phenomenon occurring across conditions variably. This is one of the reasons that the present thesis makes the distinction between a qualitative and a variable boundary strength effect. Moreover, this thesis treats coarticulation as a possible output of traditionally viewed sandhi processes, and assumes that coarticulation too can be influenced qualitatively or variably by boundary strength.

1.2.3.3 Statistical findings in support of different domain types

In the previous section we saw that inherently gradient variables, like pre-boundary lengthening and F0 scaling can be used for postulating differences of type across prosodic conditions. We discussed that the basis upon which those variables can be used as markers of constituency is when they refer to different linguistic variables, giving rise to qualitative differences. With respect to pre-boundary lengthening specifically, we discussed a possibility that lengthening might occur on different syllables across conditions. While this is a possibility, it needs to be mentioned that research has tested where and how pre-boundary lengthening occurs within languages, and it has been shown that, if pre-boundary lengthening takes place, it always occurs on the same linguistic variable across domains. Cambier-Langeveld (1997) showed that there is no evidence that different linguistic constituents are targeted for different levels in the hierarchy. Therefore, specifically final lengthening is not expected in this thesis to provide support for qualitative distinctions across domains.

Many researchers, however, have used final lengthening (and several other inherently gradient variables, like F0 scaling) to propose the existence of distinctions of type across levels, basing their proposals of *statistical* differences across conditions.

For example, we mentioned before Chavarria et al. (2004), who proposed that English has a distinction between intermediate and intonational phrases, based on the fact that there was a statistically significant difference in the amount of lengthening between the two conditions. In this thesis, such statistical significances are not reliable for postulating the existence of different types of domains. For one, it is possible that there are further gradations of strength, that an experimental design has not managed to capture. For example, Ladd & Campbell (1991) showed for English that, on the basis of pre-boundary lengthening, it is possible to identify a number of possible strength gradations, which are not necessarily differences of type. If we were to assume, however, that every statistical difference of the sort mirrors differences of type across constituents, we would be forced to postulate a much larger number of domain types, and it would be very difficult to define the linguistic basis upon which domains are constructed. Therefore, differences of statistical significance cannot be taken as reliable proof for the existence of qualitative distinctions.

In the case of a variable boundary strength effect on an gradient variable, like pre-boundary lengthening, statistical differences can also arise. This is a result like the one found in Ladd & Campbell (1991), who found lengthening across experimental conditions, and many of those were also statistically distinct. While one cannot postulate specific domain types from this finding, it is also not possible to preclude the possibility that there are differences of type across some domains, which however are acoustically manifested variably. Therefore, variable strength effects do not exclude the possibility that there are differences of type across levels, since it is possible that the experimental design has not used the phenomenon that would be appropriate for identifying differences of type. This is an added reason why it is important for researchers to specify the types of qualitative boundary strength effects expected when postulating labelled domains.

This section presented the types of acoustic evidence used in this thesis for supporting the existence of labelled domains. The use of the specific terminology and the formulation of specific expectations aims at addressing methodological issues with respect to past research, as those were discussed throughout the Introductory

section. The rest of this introduction presents the research question of this thesis, and the general methodological strategy to be followed.

1.3 This Thesis

1.3.1 Research Question

This introduction discussed how the notion of prosodic structure moved from a view entailing solely differences of boundary strength across levels to proposals assuming that it comprises of a given set of qualitatively distinct domains. It was discussed how the proposal of labelled domains making up the structure needs to be tested acoustically, and how findings pointing to differences of boundary strength were discarded in favour of proposing a given set of prosodic levels in the structure.

The main goals of the present thesis, therefore, are:

1. To test whether the acoustic manifestation of boundary strength (on segmental and suprasegmental processes) supports the existence of qualitatively distinct domain types (labelled prosodic domains).
2. To use a methodology that allows a clear overview of how prosodic constituency is mirrored in the acoustic signal, by using many processes from one language. This will ensure that we can address any existing differences of type, strength, or the concurrent existence of both. Moreover, the methodology aims at addressing methodological concerns with respect to past research within labelled views, as those were discussed in the previous sections (e.g. separating the issue of whether a sandhi phenomenon is phonological or not from the issue of what type of boundary strength effects are exercised on it).

1.3.2 Methodological strategy

The research question of the current thesis is tackled using a methodology aimed at addressing some of the limitations of previous studies on the identification of prosodic domains. The issue whether there is support for labelled domains is addressed by testing the effect of boundary strength on acoustic correlates to the structure. The main methodological contributions in answering this question are the following:

- Using *instrumental* analyses for all phenomena:

It has been mentioned repeatedly in this Introduction that labelled views are being criticized for providing proposals that were solely based on impressionistic analyses. The present thesis uses instrumental analyses for investigation of all phenomena, therefore, ensuring that any categorical findings are indeed the result of qualitative boundary strength effects.

- Using a *single* language:

Testing boundary strength effects on a single language ensures that all effects identified stem from the same prosodic structure. It has been one of the core assumptions of labelled views that all languages will have the proposed constituents. Although this is not a prerequisite for any of the theories proposed, it is suggested that such an assumption makes the theories stronger:

‘While there is no *a priori* reason that the phonology of a given languages must include all seven units, we will make the assumption here that this is the case’ (Nespor & Vogel 1986, pp. 11).

Studies using external sandhi for postulating prosodic domains have been seen to refer to different languages for different domains. However, the possibility that different languages have different structures cannot be excluded. Importantly, the use of several languages might lead to the postulation of more domains than necessary for many languages. For example, Nespor & Vogel (1986) investigate a number of languages, but there is no single language from which phenomena appear throughout the structure, in all five proposed domains above the word. For example, Modern Greek and Italian are two languages that are referred to very often, but still phenomena from Modern Greek are addressed for the postulation of only three out of five levels. Assuming that all five domains will appear in each language, however, makes it imperative that the entirety of at least one language is shown to entail all of them. Frota (1998), for example, offers one of the very few studies investigating a variety of phenomena in a single language, and proposed that both intonational and segmental processes supported the analysis proposed by prosodic phonology for European Portuguese, while her analysis of sandhi phenomena showed that their distribution across prosodic domains is not as

expected, and does not necessarily give rise to clear distinct differences of type (a finding which she explained with reference to Ladd's Compound Prosodic Domains).

- Using a variety of processes *in parallel*:

In this introduction we have seen how external sandhi and intonation patterns have been used for the postulation of labelled domains. However, different studies have mainly focused on one of the two processes as the main means of domain identification, assuming that all other cues will confirm their findings, or use each process in a limited number of domains. The present thesis aims at investigating external sandhi and intonation and timing patterns in parallel in all boundary conditions created. This ensures that most of the important processes used in past research for prosodic identification are addressed. More importantly, it is possible that different levels in the structure are associated with different phenomena. Therefore, using a variety of phenomena we will be able to address many levels, or identify the lack thereof. Finally, it is possible that one phenomenon is influenced variably by boundary strength, as final lengthening is expected to be, but that does not mean that the underlying structure is necessarily only comprised of differences of strength. By looking at intonation and external sandhi the possibility of categorical effects surfacing is covered as much as possible. However, the parallel investigation of final lengthening allows insights into the type of boundary strength differences across constructed domains, since final lengthening is well-known to mirror such boundary strength effects.

Therefore, a parallel investigation of cues to purported domains allows for an overview of the structure.

- Using a variety of *boundary strength conditions*:

Since the main research question is to identify whether changes of boundary strength give rise to categorical or variable differences in the manifestation of correlates to the prosodic structure, it is important to use a variety of boundary strength conditions, ensuring that they cover a wide range of possibilities from weak to strong boundaries. The present thesis examines a variety of processes under a variety of boundary conditions to make sure that

any categorical effects arising are indeed mirroring differences of type across constituents. In the discussion chapter we will also discuss variable boundary strength effects identified, and the types of information they provide for prosodic constituency.

These decisions are further elaborated upon in the following chapter, the one dedicated in the general experimental design and the production experiment.

1.4 Summary and Overview of thesis

In this chapter it was shown how the notion of boundary strength is important for our understanding of the nature of prosodic structure. While boundary strength is well-known to mirror the hierarchical nature of prosodic constituency, and while its acoustic manifestation in a number of studies suggests that there are many levels of possible strengths ranging on a scale from weak to strong in a gradient fashion, the theory of prosodic structure has proposed that there is a given set of labelled domains making up the structure, and that researchers need to define and identify those domains. This view is termed the *labelled view* in this thesis, because it gives rise to specific labelled domains, and it proposes that there is acoustic evidence from qualitative boundary strength effects on a variety of correlates to support the existence of labelled domains.

In the course of this Introduction it was shown how recent studies indicate that effects traditionally viewed as qualitative can actually be variable across boundary conditions when placed under the scrutiny of acoustic and articulatory measurements. These findings mirror the proposal of initial studies that boundary strength can give rise to solely differences of strength across levels. It is, therefore, proposed in this thesis to investigate a variety of processes that in the past have been used by labelled views to postulate specific labelled domains, place them under a variety of boundary strength conditions, and test acoustically whether their manifestation supports the existence of a specific set of labelled domains, as labelled views propose.

The rest of this thesis consists of six chapters; Chapter 2 - the General Experimental Design - discusses the phenomena used to tackle this issue, the way boundary

conditions were created to test the effect of boundary strength on those correlates, and the general experimental method for constructing items and performing the production experiment. The following four chapters are each dedicated to the investigation of boundary strength effects on the four selected correlates to prosodic constituency, i.e. pre- and post-boundary durations, post-nasal stop voicing, vowel hiatus resolution and intonational marking of constituents. The chapter of pre- and post-boundary durations is mainly aimed at testing whether the experimental boundary conditions created give rise to varying boundary strengths, as hoped for. Finally, Chapter 7 presents a general discussion of the findings with respect to whether they provide support for the existence of labelled views, along with a discussion on whether there is variability in prosodic structure that should be present in its representation, and the ramifications of these findings for current labelled views.

CHAPTER 2

General Experimental Design

Section 1.3 of the previous chapter presented the general strategy for addressing the research goals of the present thesis. A variety of acoustic correlates marking prosodic constituency are investigated in parallel in one language in order to identify whether their manifestation across a variety of boundary strength conditions supports the labelled view by providing clear distinctions for domain types across contexts. The present chapter discusses the exact phenomena chosen, and the construction of boundary conditions meant to elicit prosodic boundaries of varying strength. Moreover, it presents the template for constructing materials for all phenomena, and concludes with the method used for the production experiment.

2.1 General Design

The research question of the present thesis will be addressed by investigating phenomena in a single language (Modern Greek), in order to avoid the confound of different structures operating differently across languages, as could be the case in Nespor & Vogel's (1986) account. Modern Greek is a language on which there has been substantial research both on external sandhi and on intonation patterns, as well as on its prosodic grouping, and it has been investigated by many researchers subscribing to the labelled view. It is, therefore, an interesting language to use when testing the acoustic reality of the assumptions of the labelled view.

2.1.1 Acoustic Correlates

Four correlates were selected as key phenomena for investigating prosodic constituency, mainly based on the fact that they have been used in the past by researchers to postulate domains within the prosodic structure: boundary-related **durational** marking of constituency, **intonational** marking of constituency (mainly alignment and scaling of tones), and two sandhi processes, namely **post-nasal stop voicing** and **vowel hiatus resolution**. The selection of each of those processes further rests on specific motivation, which will be fully discussed in each of the introduction sections of the following four chapters, which are respectively dedicated to the analysis of boundary strength effects on each phenomenon separately. Following is a summary of the main questions we aim at answering by the use of each correlate:

- **Pre- and post-boundary durations**

This is the first phenomenon to be investigated in the thesis. Pre-boundary lengthening has been specifically referred to as a correlate of prosodic constituency working in tandem with F0 scaling in separating intermediate from intonational phrases (Nespor & Vogel 1986, Selkirk 1981, Pierrehumbert & Beckman 1988). The investigation of the particular correlate in this thesis is twofold; first, and mainly, it is used as a phenomenon that allows insights into whether the boundary strength conditions created in the experiment are of different strength. Pre-boundary durations have been shown in many studies to vary gradiently and to mirror boundary strength effects by showing longer durations higher in the structure. Conversely, the existence of pre-boundary lengthening can act as a means of confirming that boundary conditions created in a design have indeed managed to capture boundaries of varying strengths. This is the first goal that the use of this phenomenon is trying to achieve in the present thesis; given that the experimental question is whether there is support in the acoustic manifestation of boundary strength for labelled prosodic domains, it is important to establish that the boundary conditions used in the experiment are indeed of varying strengths. Investigating the duration of pre-boundary segments and finding gradually longer durations will ensure that the experimental design has fulfilled its purpose.

The second goal of the investigation of pre-boundary durations is to provide results that can be compared to results from F0 scaling to test whether differences mirroring qualitative boundary effects can arise across boundary conditions in a combination of pre-boundary lengthening and scaling.

- **Sandhi**-type phenomena have assumed a crucial role in prosodic domain identification, as was presented in the previous chapter. Research from prosodic phonology has relied heavily on the application of such phenomena when postulating prosodic domains. Their application across boundary conditions can offer clear support for qualitative distinctions across domains, since the assumption is that their application is regulated by boundary strength, and that categorical distinctions between boundaries mirror qualitative distinctions across underlying domains. This assumption is based on the assumptions that they are phonological processes and that they will be influenced categorically by boundary strength. However, it was seen in the introduction that some sandhi-type processes have recently been found to be the result of coarticulation, and to be influenced variably by boundary strength.

Therefore, the use of sandhi-type processes in this thesis assumes a crucial role, testing whether phenomena traditionally viewed as phonological are indeed acoustically realised as such, and mainly investigating the type of effects exercised on them by boundary strength.

This thesis examines two such phenomena in parallel. The reason for that is to tap into the distinction across processes that are considered phonological and those traditionally viewed as phonological but found to be the result of coarticulation; post-nasal stop voicing covers the former possibility, while vowel hiatus resolution the latter, since in recent studies vowel hiatus resolution in Modern Greek has been proposed to be the result of coarticulation. These two phenomena were selected from a range of available phenomena in Modern Greek (see e.g. Tserdanelis 2004, Arvaniti & Baltazani 2005, for a list of attested processes). Moreover, studies within labelled views have used both processes in the past to postulate specific prosodic domains.

- **Intonational marking** of prosodic constituency is the last correlate to prosodic structure to be investigated in this thesis. Intonation contours are believed to

be one of the main correlates of prosodic domains, and as such they form part of the investigation of this thesis. Different prosodic domains are expected to be signalled intonationally. The present chapter will look for cues in the acoustic signal for postulating distinct domains on the basis of intonational contours.

Finally, the most important contribution of the present thesis is that it investigates all those processes *in parallel*, under *the same* boundary strength manipulations. Therefore, it is possible to draw comparisons across phenomena. Moreover, it is possible to refer to more than one process at the same time for the postulation of either qualitative or variable distinctions across boundary strength contexts found in the data. Covering such a range of processes from one language allows for greater certainty in the conclusions drawn than using solely one process or a limited number of boundary conditions from a variety of languages.

2.1.2 Boundary conditions

Testing whether prosodic structure is comprised of a set of labelled domains, or whether acoustically there is no support for the claim, means testing whether there is acoustic support for qualitatively distinct domains. The question in the present thesis is tackled by investigating the effect varying boundary strengths have on the acoustic manifestation of phenomena traditionally used for postulating labelled domains. This by itself suggests that a set of boundary conditions needs to be created, carefully designed to elicit gradations across prosodic domains from weaker to stronger boundaries. Therefore, the manipulation of boundary strength assumes a crucial role in the experimental design, and it is the construction of those boundary conditions that is presented in this section.

Boundary strength manipulation in this thesis is meant to take place without making any assumptions regarding the prosodic constituency of the utterances involved, since the goal is to extrapolate some aspects of this constituency from the acoustic data. Therefore, boundary strength was manipulated in this thesis by adjusting the *syntax* and the *length* of the utterances to be produced. These two parameters are two of the most easily defined and well-known parameters for manipulating prosodic constituency. *Syntax* is well-known to influence prosodic constituency by

imposing constraints on the possible prosodic groupings; in example (2.1) the first two sentences are possible prosodic groupings, while the third isn't:

(2.1)

- a. George and Mary give blood. (one intonational phrase)
- b. George and Mary, give blood. (two intonational phrases)
- c. *George, and Mary give blood. (two intonational phrases)
- d. George, and Mary, give blood. (three intonational phrases)

(Shattuck-Hufnagel & Turk 1996, pp. 197)

The manipulation of the syntactic structure of the utterances used, therefore, is meant to give rise to boundary strengths of varying degrees. However, a structured way of manipulating the syntax was necessary, in order to make sure that the investigated utterances are indeed expected theoretically to yield stronger boundaries. For this reason, the syntactic manipulations used in this thesis abide by Nespor & Vogel's (1986) account for domains above the word level. Their account is used for two reasons; first because it offers a structured way of manipulating syntax and forming expectations regarding the strength of the resulting boundaries, and second because Nespor and Vogel's account has relied heavily on phenomena from Modern Greek to propose specific labelled domains. Therefore findings from this thesis will be directly comparable to their proposal. The resulting sentences also fitted the experimenter's native speaker intuitions that they might give rise to gradually stronger constituent boundaries.

Support that these specific syntactic manipulations can give rise to stronger boundaries came from Baltazani's (2006) paper (which was published right after the recordings for the production experiment for this thesis were finished). Three of her syntactic manipulations were identical to three out of six experimental conditions created in this thesis. The syntactic manipulations she used were meant to give rise to stronger boundaries (although in her analysis they were also assumed to give rise to different types of domains, as expected by labelled views).

The *length* of the utterances for some conditions was also manipulated in an effort to create stronger boundaries. The term *length* refers to the number of syllables making up an utterance. This decision is based on a number of studies having

reported that length is a significant structural property of language influencing the way prosodic grouping takes place, as seen in the previous chapter (Grosjean & Collins 1979, Gee & Grosjean 1983). More recently, Watson & Gibson (2004) reported that the likelihood of an intonational boundary occurring before and after each word is largely dependent on the length of the preceding and following utterances. The importance of length in prosodic constituency was also acknowledged in research postulating the indirect syntax prosody mapping. Nespor & Vogel (1986) proposed a *restructuring* rule to account for differences in constituent grouping when the length of the utterances was longer (we will return to their predictions in the Discussion chapter, where we discuss the types of effects length produced with respect to the underlying relationships across prosodic domains).

Six conditions were created by manipulating the syntax and length of the utterances; five were the result of syntactic manipulations à la Nespor and Vogel, and one of length. Nespor and Vogel's prosodic structure above the orthographic word level entails the following domains: the clitic group (CG), the phonological phrase (φ), the intonational phrase (IP), and the utterance (Utt). In the present design there are utterances for each of those levels.

Condition One (BC1)

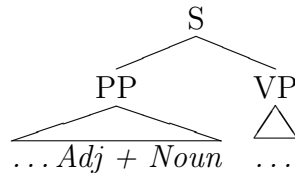
The boundary conditions constructed in this section use Nespor and Vogel's proposal as a means of manipulating syntax in a principled manner to create stronger boundary strength conditions. The preboundary word is always an Adjective, and the postboundary word a Noun, because those words easily fitted syntactic constructions amenable to Nespor and Vogel's proposal (with the exception of BC0, a condition presented at the end, since it involves different words representing a lower boundary not captured by an adjective+noun boundary).

The first condition (BC1) is based on Nespor and Vogel's Clitic Group¹, which is formed by combining all ω s that are independent words and any clitics on the left hand side of the ω ; for example $\mu\omicron\upsilon \tau\omicron \acute{\epsilon}\delta\omega\sigma\epsilon$ [mu to 'eðose] '(he) gave it (to)

¹Immediately higher than the phonological word ω , which is comprised of the stem and any prefixes or affixes to the word. For example, the word $\sigma\upsilon\nu\acute{\alpha}\nu\theta\rho\omega\pi\omicron\varsigma$ in Modern Greek [si'nanθropos] 'fellow being' < sin 'plus' + ['anθropos] 'human' forms one ω .

me' (Nespor & Vogel 1986, pp. 155) forms one clitic group. BC1 in the present experiment consists of a boundary falling between two clitic groups. The Adjective and the Noun before and after the boundary form clitic groups on their own, since they are part of the same syntactic phrase, in this case a prepositional phrase. Example (2.2a) shows a sentence for this condition with its syntactic analysis, while example (2.2b) shows its prosodic analysis (only the relevant words are marked up).

For each correlate investigated 6 to 8 items were constructed; for example for the investigation of pre-boundary lengthening, eight items were constructed and then embedded in sentences for each boundary condition (a summary of the number of sentences constructed can be found in Table 2.1 further down the chapter).



(2.2)

- a. [a'po a'ftin tin **ci'notopi ko'pela**]_{PP}, [to pe'rimena]_{VP}.
*from this the common² girl*_{SING-ACC}, *it expected*_{IMPERF-1s}
 'From this common girl, I expected it.'
- b. a'po a'ftin [_Ctin ci'notopi]_C [_Cko'pela]_C to pe'rimena

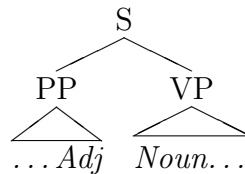
The length of the sentences was kept relatively constant (ranging from 16-19 syllables). The position (in terms of number of syllables) of the pre- and post-boundary words did not change relative to the onset and the offset of the sentence .

Condition Two (BC2)

The second boundary condition (BC2) entails again an adjective followed by a noun, but this time the two words belong to different syntactic phrases within the same clause; the pre-boundary word is the part of a prepositional phrase, and the post-boundary word is a noun operating as the object of the verb of the main clause.

²The word *κωνότοπη* denotes a person who believes in clichés, and is generally boring. When referring to someone's appearance again it denotes a boring appearance. It is a mixture of boring and ordinary.

This manipulation makes them have different lexical heads, and therefore in Nespor and Vogel's terms they belong to different phonological phrases. Example (2.3a) shows the syntactic construction, and (2.3b) the anticipated prosodic structure.



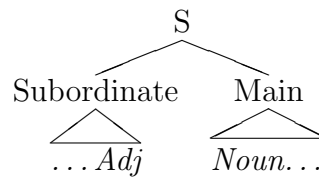
(2.3)

- a. [me em'fanisi **ci'notopi**]_{PP} [ko'pela ðen θa 'vri]_{VP}.
*with appearance common*_{SING-ACC} *girl* _{SING-ACC} *not* _{NEG} *will*_{aux} *find*_{FUT-3s}
 With this common appearance he will not find a girl.
- b. [_{IP} [φ me em'fanisi ci'notopi]φ [_{IP} [φ ko'pela ðen θa 'vri]φ]_{IP}]

The length of the sentence was again relatively constant across sentences, ranging from 16-19 syllables throughout all materials, as was the relative position of the materials within the sentences.

Condition Three (BC3)

Again aiming to address a higher order prosodic domain, the pre- and post-boundary words are syntactically manipulated as to entail a disjuncture at the intonational phrase level. This is done by placing them in two different clauses (a main and a subordinate clause), which are not, however, structurally directly attached to the highest node of the structure, occupied by the utterance level. While the sentences are kept to roughly the same length (ranging between 17 and 21 syllables), the pre- and post-boundary words belong to different clauses. The Adjective is the last word of a subordinate Clause, while the Noun is the first word of a main clause, starting with a Noun Phrase (2.4). This type of constructions corresponds to a prosodic boundary of Intonational Phrases in Nespor and Vogel's analysis. Example (2.4a) shows the syntactic construction of the sentences, and (2.4b) the prosodic.



(2.4)

- a. [a'fu 'eçis em'fanisi **ci'notopi**]_{Sub}, [ko'pela ðen θa 'vris]_{Main}.
since have_{PRES-2s} appearance common_{FEM-ACC}, girl_{FEM-ACC} not_{NEG}
will_{aux} find_{FUT-2s}
 ‘Since you have a common appearance, you will not find a girl.’
- b. [_{IP}a'fu 'eçis em'fanisi ci'notopi]_{IP} [_{IP}ko'pela ðen θa 'vris]_{IP}

While conditions BC2 and BC3 are fairly similar in their segmental make-up, their syntactic construction is very different giving rise to different expectations. Moreover, while between the adjective and nouns in the written form of BC3 there is a comma separating the main from the subordinate clause, there is no such marker between words in BC2.

Condition Four (BC4)

The next level in Nespor and Vogel’s analysis is that of the Utterance. However, before making a construction mirroring that level, it was important to ensure that not only the syntactic structure is mirrored in the data, since any effects found could be argued to be based on a direct syntax effect. Therefore, the *length* of the utterances is manipulated in this domain, instead of their syntax. The same syntactic construction is used as for BC3, but both the main and the subordinate clauses were longer. This was done by adding prepositional phrases in the subordinate and main clauses, therefore the length of the main clause varied between 15 and 18 syllables, and the subordinate from 16 to 18 syllables, with the total length of the utterances ranging from 34 to 36 syllables (see examples (2.5a and b) for the syntactic and prosodic constructions respectively).

- (2.5) [a'fu 'eçis ka'ta jeni'ci omolo'jia em'fanisi **ci'notopi**]_{Sub}, [ko'pela ðen θa 'vris xo'ris rizi'ci ala'ji]_{Main}.
since have_{PRES-2s} by general consent_{FEM-ACC} appearance

*common*_{FEM-ACC}, *girl*_{FEM-ACC} *not*_{NEG} *will*_{aux} *find*_{FUT-2s} *without radical change*_{FEM-ACC}

‘Since everyone agrees that you have a common appearance, you will not find a girl without a radical change.’

- b. [IPa'fu 'eçis ka'ta jeni'ci omolo'jia em'fanisi ci'notopi]_{IP} [IPko'pela ðen θa 'vris xo'ris rizi'ci ala'ji]_{IP}

Condition Five (BC5)

This is expected to be the strongest disjuncture between the adjective and noun. The same sentences created for BC4 were used, but instead of having a main and a subordinate clause separated by commas, two main clauses were created. Thus, there was a full-stop between pre- and post-boundary words, as can be seen in example(2.6). Example (2.6b) shows that the two sentences form two utterances in Nespor and Vogel’s analysis.

(2.6)

- a. ['eçis ka'ta jeni'ci omolo'jia em'fanisi **ci'notopi**]_{Main}. [**ko'pela** ðen θa 'vris xo'ris rizi'ci ala'ji]_{Main}.

*have*_{PRES-2s} *by general consent*_{FEM-ACC} *appearance common*_{FEM-ACC} - *girl*_{FEM-ACC} *not*_{NEG} *will find*_{FUT-2s} *without radical change*_{FEM-ACC}

Everyone agrees that you have a common appearance. You will not find a girl without radical change.

- b. [Utt'eçis ka'ta jeni'ki omolo'jia em'fanisi ci'notopi]_{Utt} [Uttko'pela ðen θa 'vris xo'ris rizi'ci ala'ji]_{Utt}

This is a syntactic manipulation which kept the length of the utterances before and after the boundary the same, as in BC4.

Therefore, it was shown in this section how the syntax and the length of the utterances involved were manipulated with the goal of creating gradually stronger boundary conditions. It is hypothesized that the boundary strength across conditions BC1 and BC5 gets gradually stronger (on the basis of Nespor and Vogel’s predictions), and the prediction from the labelled view is that each of those domains is qualitatively different from the others. As a starting point, Nespor &

Vogel's (1986) analysis was used for providing a principled manner of manipulating boundary strength, and forming expectations as to why each domain constructed might give rise to gradually stronger boundaries.

Extra Condition (BC0) with different template

The prosodic condition immediately lower to the clitic group in Nespor and Vogel's hierarchy is that of the phonological word. For a phonological word to distinguish from a clitic group, a clitic needs to attach to the ω ; for example, the word [ko'pela] forms a phonological word on its own, but given that there are no clitics attached to it, it also forms a clitic group. However, if it is preceded by a definite article, e.g. [tin ko'pela], then the two together form a clitic group. In order to create a condition lower to the clitic group (which was our BC1), the template Adjective+Noun for the pre- and post-boundary words does not work. Therefore, the Nouns used in post-boundary positions were kept, but the Adjectives were replaced by definite articles³. For example, in the case of [ki'notopi ko'pela], the resulting item would be [tin ko'pela]. Therefore, this last condition is expected to be the weakest with a juncture of a phonological word; between the definite article and the noun the prosodic grouping is always expected to look like [C [ω tin] ω [ω ko'pela] ω]C. The position of the words within the sentences was again kept constant, as was the number of syllables across items, as much as possible.

The Adjective (or definite article) preceding the boundary, and the Noun following it, provided the template for the syntactic constructions. Stimuli were constructed for each phenomenon and were then embedded in utterances constructed using the six boundary conditions just described.

2.1.3 Stimulus design

As already mentioned, the effect of prosodic boundary strength was investigated by changing the strength of this boundary between two words. The pre-boundary word was always an ADJECTIVE, and the post-boundary word a NOUN.

³For the materials in post-nasal stop voicing definite articles, negative particles and subordinating particles were also used, but see Chapter 4 for an analysis of the materials for this phenomenon.

Both pre- and post-boundary words were almost always tetrasyllabic, with open CV syllables, if possible (Example 2.7). The position of the stress was kept constant for all items; the pre-boundary word was stressed on the antepenultimate syllable, and the post-boundary word on the second syllable after the boundary. The reason for that is that stress has been found to influence the realisation of segments in numerous ways, e.g. changing their articulatory patterns, affecting their duration, and changing the F0 pattern of the words involved (Beckman & Edwards 1994, Arvaniti 2000, Turk & White 1999, Turk & Shattuck-Hufnagel 2000, among others). Furthermore, it is important for the purposes of this experiment that the stressed syllables before and after the boundary are not close together, given that we would like to avoid phonetic pressure potentially leading to truncation of F0 (Arvaniti, Ladd & Mennen 1998, Prieto, d' Imperio & Gili Fivela 2005). Hence, the number of syllables and position of the stress on the pre- and post-boundary words were carefully selected. Example 2.7 shows the general template for constructing all items, and an example taken from the items constructed to measure pre- and post-boundary durations. In this template 'σ' stands for syllable, # is the boundary sign, and ' stands for the stress on the following syllable.

(2.7)	ADJECTIVE	NOUN
	Template: σ 'σ σ σ # σ 'σ σ (σ)	
	Example: ci 'no to pi # ko 'pe la	
	ordinary/common	girl

This template was used to create items for investigating each of the four phenomena analysed in this thesis. Three sets of materials were constructed; one for measuring pre- and post-boundary durations, one for vowel hiatus resolution, and one for post-nasal stop voicing.

Stop voicing and vowel hiatus require their own set of items, since they involve different segmental make-ups giving rise to the sandhi-type process. Stop voicing occurs when a voiceless stop is preceded by a nasal. Therefore, the pre-boundary segment needs to be a nasal and the post boundary segment a voiceless stop. Vowel hiatus, on the other hand, occurs when two vowels are adjacent to each other, in

which case the pre- and post-boundary segments need to be vowels. Therefore, by definition, different sets are required.

For measuring pre-boundary durations, one could propose that the same datasets, as for stop voicing and vowel hiatus, could be used. This would ensure that most conclusions are drawn from the exact same renditions. However, the datasets from stop voicing and vowel hiatus are not optimal for measuring pre- and post-boundary durations. In the case of post-nasal stop voicing, it has been suggested that the pre-boundary nasal might get deleted, influencing any duration patterns in pre-boundary positions (Arvaniti 1999, Arvaniti & Joseph 1999). In the case of vowel hiatus, there is no clear landmark to mark the position of the boundary within the vowel pair continuum (for example if the boundary falls between vowels [a#o], it is not possible to mark the exact place of the boundary), therefore it is not possible to separate pre- from post-boundary effects. Consequently, a different set of materials is constructed for measuring boundary related durations.

Only for the investigation of intonation contours did we use the same materials as for pre-boundary durations and post-nasal stop voicing, since it is desired that as many results as possible come from the same dataset. The data from vowel hiatus resolution were not suitable for analysing F0 again because of the difficulty in identifying where the boundary is located within V1V2. Since we were measuring the alignment of the pre-boundary peak with respect to the boundary, knowing where the boundary is located is very important. The way items were constructed for each phenomenon separately (using template 2.7) is presented in the Methodology section of each relevant chapter.

2.2 Method

2.2.1 Participants

Five native speakers of Standard Modern Greek were recorded producing the materials; two male (MS and CP) and three female (OM, AT, and SM). They were all in their twenties at the time of the recordings and were all students of the University of Edinburgh, having lived in the UK for a period of time no longer than a year. They were all brought up in Athens, none reported a speech, hearing, or reading

problem, they were all naive as to the purpose of the experiment and none had a linguistic background. They were paid for their participation.

2.2.2 Procedure and recordings

A carefully planned production experiment was designed, within which speakers were asked to read out sentences containing the items and boundary conditions created to test the general research question. In order to ensure that their productions do not take the form of a reading list of sentences, and are as natural as possible, participants were prompted with a **question** from the experimenter for each sentence they had to produce. The reason for that was threefold; firstly, the participant was prompted with the words they were about to produce, and was given the time to become familiar with the sentence. Secondly, the question asked for the focus to be placed on a particular part of the sentence (at the final verb, placed at the end of each sentence, therefore not on the carrier words), so that all participants produced the same renditions for all sentences and the items under analyses were not focused, which is known to alter the phonetic manifestation of the relevant sounds (Baltazani & Jun 1999). Finally, question-answering made the recording sessions shorter than, for example, embedding the materials in paragraphs (which is a different method for eliciting more natural productions), therefore making it possible to record more items and repetitions. Example (2.8) shows an instance of question-answer pair along with the translation (the underlined words are the ones receiving focus and the words in bold are the carrier words).

(2.8) Question: [a'po a'ftin tin **ci'notopi ko'pela** to pe'rimeses]?

Translation: From this common girl, did you expect it?

Answer: [a'po a'ftin tin **ci'notopi ko'pela** to pe'rimesa]

Translation: From this common girl, I expected it.

On two occasions in this thesis a perception experiment was coupled with the production experiment (namely in the two chapters investigating the application of sandhi-type phenomena), but the motivation and design of each of those is specifically presented in the relevant chapters.

Controlling for possible speech rate effects was addressed using randomised lists; each speaker had a list of the sentences randomised differently, so if there were any effects from being slower at the beginning and speeding up in the end, these were not localised on the utterances of say BC5, which would be the last utterances to be produced should the utterances not be randomised (thus also having the same effect across all speakers). Following the recordings, possible pertaining speech rate effects were checked by measuring the duration of a ‘control’ word, which appeared in sentences for all boundary conditions, and which was measured to make sure the rate of speech did not vary across recordings (see Chapter 3 for this analysis, which supported the claim that there were no speech rate effects for any of the speakers).

In total approximately 2,300 sentences were recorded from all five speakers for all phenomena (table 2.1 shows how many sentences were produced for each correlate, and how many in total). The items constructed for measuring boundary related durations (named “Durations” in the table) refer to the items specifically recorded to measure pre- and post-boundary durations. While these were items specifically designed to investigate the effect of boundary strength on duration patterns, they are not the only durational measurements performed in this thesis, since the duration of all segments in all phenomena are also measured. This specific dataset, however, is designed to address the issue of boundary related durations, and therefore includes items with appropriate segmental make-up.

Finally, this table is broken down into two panels; the top one presents items for conditions BC1-BC5, and the bottom one for BC0. The reason for this is that BC0 is a different construction (recall the difference in construction: particle + verb vs adjective+noun), and has fewer items⁴.

During the recordings, the participants were asked to read in a casual manner, as they would do if they were speaking to someone. They were asked to maintain

⁴BC0 items were initially constructed only for stop-voicing, since it is the only process which (according to my native speaker’s intuitions and Nespor and Vogel’s analysis - see Chapter 4) would need a lower domain for the process to occur. Therefore, all speakers were recorded producing BC0 items for post-nasal stop voicing. However, while the recordings were underway, it became clear that for comparison reasons it would be good to have BC0 items for all phenomena investigated. Therefore, extra materials were added, but only two speakers were still in the recording process. Therefore, for BC0 items of pre-boundary lengthening and vowel hiatus we only have production data from the two male speakers, and for stop voicing we have production data from all speakers.

Process		Carrier Wds	Rep	BoundCond	Speakers	TOTAL
Boundary Conditions BC1-BC5						
Durations		8	2	5	5	400
Vowel Hiatus	2 [ao],[oa]	6	2	5	5	600
	2 [aa],[oo] ^a	6	2	5	5	600
Stop-Voicing	[p,t,k]	8	2	5	5	400
	[b,d,g] ^a	3	2	5	5	150
F0	used materials from Durations and Stop Voicing					(-)
Boundary Condition BC0						
Durations		6	2	1	2	24
Vowel Hiatus	2 [ao],[oa]	6	2	1	2	24
	2 [aa],[oo] ^a	6	2	1	2	24
Stop-Voicing		15	2	1	5	150
TOTAL						2,272

^abaseline items

Table 2.1: Exact numbers of sentences recorded for all speakers for each of the processes.

the same distance from the microphone throughout the recording by not moving as much as possible, and not to turn pages while speaking to avoid overlapping noise with speech. If the participant made a disfluency or placed the focus on the wrong place, they were asked to repeat the sentence and were prompted with the question again.

Recordings took place at the recording booth of the Linguistics and English Language Department of the University of Edinburgh using an AKG CK 98 Hypercardoid condenser Microphone with a MOTU 828 Mk2 interface. The software used for the recording was the Sonar 4 Studio Edition; 48KHz 16bit WAV files were saved straight to the hard drive.

2.2.3 Analysis

Three datasets were created from the recordings; one containing materials for boundary related durations, one for post-nasal stop voicing and one for vowel hiatus resolution (recall we will be using the first two datasets for the analysis of intonational contours). The way materials within each dataset were segmented for the purposes of the analysis of each phenomenon is presented in chapters 3, 4, 5,

and 6 respectively. In general, the sentences produced by the participants were extracted from the recordings in a semi-automatic way (marking the beginning and end of each sentence and using a script to extract them and label them⁵). Then each sentence was segmented (using the segmentation techniques we will present in each chapter) and was labelled using a Break Index analysis, which was an intuitive way of making sure the same type of disjuncture existed across processes, phenomena and speakers, and mainly to make sure that any utterances with wrong focus, or unnatural intonation, were discarded. This Break Index analysis was not used in any other way, and was not meant to represent any type of prosodic domain identification.

After the segmentation of the relevant landmarks took place, all measurements were automatically extracted from the waveforms and inserted to Excel for post-processing. SPSS was used for most statistical analyses (with the exception of discriminant analyses which sometimes were performed using R).

2.3 Summary

In this chapter the general experimental design for addressing the research question was presented. Four phenomena have been selected as representative correlates of prosodic constituency, and six boundary conditions are created by manipulating the syntax and the length of utterances in order to elicit gradually stronger prosodic boundaries. The motivation and way the syntax and length of the utterances were manipulated was presented, along with the template for the construction of all items across phenomena.

The following four chapters are dedicated to the investigation of boundary strength effects on the phonetic realisation of the four correlates. First the effect on boundary-related durations is investigated in the following chapter with the goal of identifying whether the experimental conditions created capture the hierarchical nature of the prosodic structure.

⁵All scripts used in this thesis are written by the experimenter unless specified otherwise, and can be shared with anyone interested in using them.

CHAPTER 3

Pre- and post-boundary durations

Durational patterns have been extensively referred to in past research to support the existence of a hierarchically layered prosodic structure, in that longer durations are generally found in pre-boundary positions close to stronger prosodic boundaries. The present chapter investigates the duration of pre- and post-boundary segments in several boundary conditions created for this thesis to ensure and confirm that the experimental design presented in Chapter 2 captures a variety of boundary strengths, mirroring the hierarchical nature of the structure. An experiment is presented testing pre- and post-boundary durations in six boundary conditions, showing that in pre-boundary positions the conditions created for this thesis do indeed give rise to gradually longer durations higher in the structure. The findings are also discussed with respect to how far the phenomenon spreads before the boundary, and with respect to past research in Modern Greek.

3.1 Introduction

3.1.1 Durations as correlates to prosodic structure

The lengthening of segments in some contexts is an effect noted early on in research on prosody. One explanation offered by past research has been that pre-boundary lengthening is directly regulated by syntactic constituency, in that longer durations can be found in the vicinity of stronger *syntactic* boundaries (Klatt 1975, Cooper & Paccia-Cooper 1980). Pre-boundary lengthening was one of the crucial correlates to give rise to the notion of *boundary strength* presented in Cooper & Sorensen (1981),

as was discussed in Chapter 1. However, many examples of syntactic constructions not yielding the anticipated lengthenings (see Shattuck-Hufnagel & Turk 1996, for relevant examples) have shown that pre-boundary lengthening patterns are regulated by the *prosodic constituency*, not the syntactic one.

Boundary-related timing patterns are viewed as an important correlate of prosodic structure (Wightman et al. 1992, Price et al. 1991). The present chapter investigates the duration of segments before and after a variety of boundary contexts, and its main goal is to test whether the six designed boundary conditions have managed to create gradually stronger boundaries. These six boundary conditions are designed to elicit stronger boundaries, and to represent levels of strength that are as ‘adjacent’ as possible. That is, the strength differences from one condition to the other will be similar; for example, it is not desired to compare conditions like the prosodic word and the Utterance (in Nespor and Vogel’s terms), since those are by definition expected to be markedly different. Rather, it is desired to test conditions whose strength does not change markedly, so as to ensure that any qualitative boundary strength effects found mirror differences of type across constituents.

The experimental design, as presented in the previous chapter, was carefully designed to address this gradual strengthening of boundary strength across boundary conditions. By testing the duration of pre- and post-boundary segments the goal is to confirm that the experimental design has succeeded in capturing a strengthening of boundaries from BC0 to BC5.

3.1.1.1 Pre-boundary durations

Durations of pre-boundary segments are known to undergo lengthening when in the vicinity of stronger prosodic boundaries, a phenomenon called **pre-boundary lengthening**. This is a phenomenon that has been attested in many languages, like English (Wightman et al. 1992, Price et al. 1991, Turk & Shattuck-Hufnagel 2007), Korean (Cho & Keating 2001), Dutch (Cambier-Langeveld 1997), Finnish (Nakai, Kunnari, Turk, Suomi & Ylitalo 2008), Hungarian (Hockey & Fagyal 1999), to name a few.

Pre-boundary durations are thought to mirror the nature of prosodic structure, and in particular the fact that prosodic boundaries higher in the structure are *stronger* than those lower in the structure. Several findings in past research have highlighted the link between timing patterns and boundary strength; Wightman et al. (1992) investigated how prosodic constituency is mirrored in the duration of segments for domains that are perceived to be of varying boundary strengths, and reported that duration of the sounds before the prosodic boundary was found to be gradually longer higher in the structure (see Chapter 1 for a review of how lengthening patterns have been linked to boundary strength). The present chapter relies on the capacity of pre-boundary lengthening to mirror strength differences, and uses the phenomenon as a methodological tool for confirming that the experimental design has captured differences of strength across the constructed boundary conditions.

3.1.1.2 Post-boundary durations

More recent research has shown that prosodic constituency also influences post-boundary segments. With respect to timing patterns, the durational changes that segments undergo on the right hand side of the boundary are part of a phenomenon referred to as *initial strengthening*, which itself is part of ‘prosodic strengthening’, the ‘temporal and/or spatial expansion of articulation due to accent and/or prosodic boundary’ (Cho 2005). In this chapter, we focus on prosodic boundary effects on duration patterns, but prosodic strengthening has been found to affect a number of articulatory and phonetic correlates. For example, it has been linked with a higher rate of glottalized vowels in higher order prosodic domains (Dilley, Shattuck-Hufnagel & Ostendorf 1996), with longer durations for alveolar stops in English (Pierrehumbert & Talkin 1992), with more linguopalatal contact during the production of voiceless stops in higher order domains in Korean (Cho & Keating 2001), and in one or the other way it has been attested in a number of languages, like English (Fougeron & Keating 1997, Cho 2004, Cho 2005, Cho, McQueen & Cox 2007), French (Fougeron 2001, Tabain 2003), Dutch (Cho & McQueen 2005), Korean (Cho & Jun 2000, Keating 2003), Taiwanese (Hsu & Jun 1998, Hayashi et al. 1999, Keating 2003).

While in pre-boundary positions boundary strength always leads to the lengthening of the relevant segments, in post-boundary positions the effect has been shown to vary cross-linguistically. For many languages initial strengthening has been found to lengthen the duration of the segments involved; however, recent research suggested that prosodically driven acoustic characteristics are bound by language specific constraints. For example Cho & McQueen (2005), in testing the perceptual effect of duration differences across prosodic conditions in Dutch, found that - contrary to English, where VOT of voiceless stops gets lengthened in initial positions of higher domains - shorter VOTs were used for /t/s in higher order domains. It is, therefore, interesting to investigate more languages in order to understand the types of effects boundary strength exercises on segments adjacent to the boundary.

The pre- and post-boundary durational behaviour of segments is, therefore, viewed as a clear correlate of boundary strength. As such it is used in the present chapter; pre-boundary durations are investigated with the goal of ensuring that the experimental design has captured a variety of boundary strengths.

3.1.2 Durations and prosodic constituency identification in labelled views

Pre-boundary durations have sometimes been used by past research to offer support for the postulation of specific domains within labelled views. For example, Selkirk (1978) refers to ‘pre-pausal lengthening’ making the prediction that it signals Intonational Phrases:

“It is between intonational phrases (and only between them, we would claim) that one finds potential pauses. And it is at the end of I[ntonational phrases] that so-called pre-pausal lengthening is found” (Selkirk 1978, pp. 135)

She clearly suggests a categorical distinction between the highest domains in her prosodic structure and the lower domains, according to which lengthening can only take place across intonational phrases. The intonational phrase is also believed

to be marked by substantial pre-boundary lengthening by Pierrehumbert & Beckman (1988), and both analyses make the prediction of a phonetically categorical distinction.

Beckman & Edwards (1990), investigating differences in pre-boundary lengthening between levels of prosodic constituents (among other things), have shown that in their data at least a binary distinction arose between word-final and phrase-final lengthening. However, they are careful in not proposing that the categorical distinction they refer to is derived only from the durational differences they found, but also suggest that there are other factors supporting the distinction, like intonation contours and consistent findings across speakers and rates of speech.

These studies made use of a binary distinction between lowest and higher domains of prosodic structure in terms of pre-boundary lengthening to propose two distinct levels of prosodic constituency. It has been suggested, though, by several studies (as it was shown in Chapter 1) that a binary distinction between two extreme domains from the prosodic structure does not necessarily capture the whole range of possible boundary strength differences that can be distinguished based on lengthening patterns. Wightman et al. (1992) specifically set out to investigate differences in pre-boundary lengthening between boundary strength conditions lying in-between the lowest and highest end points of the continuum. They mention that:

“It is not unreasonable to suspect that duration cues may be able to distinguish among different level boundaries within [phrase-final versus phrase-internal] groups.” (Wightman et al. 1992, pp. 1710)

And, indeed, all of Wightman et al. (1992), Price et al. (1991) and Ladd & Campbell (1991) found many gradations of boundary strength mirrored in the statistical differences across conditions of pre-boundary lengthening, highlighting the fact that the phenomenon can be used for specifying the hierarchical nature of constituency, but not for the specific identification of prosodic domains, since there are no clear landmarks for categorical distinctions.

Given the difficulty pre-boundary durations pose for domain identification, research within labelled views has tried to make use of the phenomenon *in tandem* with other

correlates to prosodic structure, to enhance or provide further support of the existence of distinct labelled prosodic boundaries. For example, Frota (1998) used lengthening patterns along with the application of sandhi processes and intonation contours to propose the existence of the phonological phrase for European Portuguese, while Pierrehumbert & Beckman (1988) treated pre-boundary lengthening as a cue working parallel to intonational marking in separating different types of prosodic domains. They suggested that the intonational marking of intermediate phrases might not be tonally different to the immediately lower level, but that its categorical difference probably arises from the scaling of the tones and the duration of the segments involved, in the sense that higher scaling and longer durations will signal the different domain.

However, given that there is no clear categorical acoustic reason to distinguish renditions that vary in a continuum of possible values, in the present thesis we take the stand that statistical differences across a continuum of variables cannot act as an identification point for categorical differences. By that it is meant that, on the basis of gradient boundary strength effects, like those found on pre-boundary lengthening, it is not possible to propose qualitative differences of type across constituents.

Finally, it is often the case that, even in research assuming the existence of labelled domains, the anticipated statistical differences separating the domains are not always found. For example, in recent years there has been a number of studies investigating phonetic and articulatory correlates of prosodic structure, with the goal of understanding the underlying articulatory mechanisms and strategies involved in signalling prosodic structure (e.g. Fougeron & Keating 1997, Fougeron 2001, Keating 2003, Cho & Keating 2001, among many). These studies most commonly use sentences labelled on the basis of the intonational approach, that is mainly distinguishing three distinct prosodic levels, expecting to test how these are signalled acoustically and articulatorily. While the assumption is that the domains will be categorically distinct, results on durations from all these studies have shown that the differences between purported domains are not categorical. For example, Fougeron (2001) and Cho & Keating (2001) investigated the duration of post-boundary segments in Korean and French respectively, and, while the prosodic

coding they used for classifying different boundary conditions relied upon the intonational approach (Jun's (1993) analysis of Korean entailing Words, Accentual Phrases and Intonational Phrases, with an added domain called Utterance, and for French an analysis entailing Syllables, Words, Accentual Phrases and Intonational Phrases), both studies reported that the duration of post-boundary segments did not always statistically differentiate the expected categorical differences across domains. Additionally, most of those studies report on large variability in the types of groupings signalled, and in the distribution of values within prosodic domains. Such findings highlight the issue that prosodic constituency does not seem to be easily identified on the basis of such inherently gradient variables, which cannot provide solid phonological support for the postulation of domains.

To conclude, the present thesis only looks for qualitative boundary strength effects of the type discussed in Chapter 1. Statistical differences across domains on inherently gradient variables are not taken to mirror qualitative distinctions of type across prosodic domains. Given that previous research has established that pre-boundary lengthening undergoes variable boundary strength effects, and that these effects do not change qualitatively across domains, this phenomenon is only used as a methodological tool for identifying whether the constructed boundary conditions give rise to varying boundary strengths, and not for identifying specific domains.

3.1.3 Boundary-related duration patterns in Modern Greek

A variety of effects have been reported to influence timing patterns in Modern Greek; stress (Arvaniti 2000), accent and focus (Baltazani & Jun 1999), the position of the segment in the word (Baltazani 2007a) etc. have all been reported to influence the duration of segments. Research on effects of *prosodic phrasing* on durational patterns has been reported only for the lower levels of Modern Greek's prosodic structure. It has been suggested that Modern Greek does not exhibit final lengthening at the juncture of prosodic words. Katsika (2007) tested the duration of the preboundary vowel and nasal in sequences like [to 'nano] vs [ton 'ano], and found no durational differences between the two conditions for the /n/, suggesting that prosodic constituency was not signalled durationally at all between those domains. The duration of the first postboundary vowel in her data, however, was

found to be greater word-initially, supporting a view of initial strengthening even at this low domain (although the durational differences were very small). Similarly Arvaniti (2000) (and Botinis 1989, in a similar experiment with similar findings) presented evidence that Modern Greek does not exhibit lengthening across prosodic words, and that initial lengthening might be apparent. In an experiment that was testing the phonetics of stress in Greek she investigated the duration of words like [ˈpapa] vs [paˈpa] and reported that the final segment of the words was not lengthened in comparison to the word medial segment, therefore not exhibiting prosodic word final lengthening, and word initial consonants were longer than word medial ones (an effect that disappeared on the following vowel/syllable), suggesting initial strengthening reaching only up to the first segment of the postboundary word (for somewhat different results, but with also different prosodic constructions, see Tserdanelis 2004).

There are only three studies on Modern Greek reporting on durations from analyses investigating higher order domains. Baltazani (2006), while investigating the effect of prosodic phrasing and focus on vowel hiatus resolution, reported that the duration of the vowel pairs across prosodic words were significantly shorter than those of higher order domains, thus proposing that phrase-final lengthening does exist in Modern Greek, and concluded that it signals a categorical difference between the lower and higher domains. She also specifically investigated the effect of prosodic structure on pre-boundary durations in Baltazani (2007*b*), where again she reported a categorical difference between prosodic words and intermediate phrases. Additionally, by testing differences between word initial, medial, and final syllables, she supported previous findings that Modern Greek does not exhibit word final lengthening. Finally, Kainada (2007)¹, investigating vowel hiatus resolution, reported on longer durations in higher order domains for vowel pairs, and her data mirrored a hierarchical gradient effect with longer durations in higher order domains.

There is a difference between Baltazani's (2006) and Kainada's (2007) findings, in that Baltazani reported a categorical difference between lower and higher domains, while Kainada found gradiently varying effects with adjacent conditions not being

¹This is a set from the data from the current thesis presented at the 16th International Congress of Phonetic Sciences.

significantly different. This difference is crucial to the current discussion; given the inherently gradient nature of duration, a categorical difference across domains is taken generally to be one that gives rise to a statistical difference, as in the case of Baltazani's data. However, it might be the case that the specific experimental design has not addressed conditions that could give rise to intermediate gradations of lengthening (like conditions with stronger or weaker boundaries), which seems to be the case in Kainada's data. This supports the point of view taken here that statistical differences in gradient variables cannot offer reliable support for qualitative distinctions of type across domains. Therefore, while categorical differences can be found across prosodic contexts, in the sense that they arise as statistical differences, this does not support the conclusion that they mirror underlying qualitative differences of type across prosodic domains. This is an important distinction that is maintained in the present thesis.

The present chapter aims at adding to the existing literature on timing patterns at the vicinity of prosodic boundaries above the word level for Modern Greek, that is levels that have not been investigated as such. This investigation will provide data with which the results of future research can be compared. Moreover, results from Chapter 6 on the scaling of F0 targets across boundary conditions will be linked to results from this chapter in an integrated analysis of boundary strength effects on correlates to prosodic structure. Finally, the present investigation on Modern Greek aims at adding cross-linguistic evidence to the issue of 'how far' the effect of boundary strength spreads before the prosodic boundary, a question to which we now turn in the following section.

3.1.4 How far does pre-boundary lengthening spread?

Since this chapter constitutes the first effort to directly address the issue of pre- and post-boundary duration patterns for Modern Greek above the word level, one further issue will be discussed, i.e. how far pre-boundary lengthening spreads. This is a secondary goal of the present chapter, and aims to add to the existing literature on the matter by providing cross-linguistic data. Recent research has reported that it can occur on pre-boundary segments, syllables (Byrd & Saltzman

1998), or more extended chunks of speech (e.g. Wightman et al. 1992, Cambier-Langeveld 1997, Turk & Shattuck-Hufnagel 2007, Nakai et al. 2008). The exact locus of lengthening is therefore still unclear; while some researchers suggest that it is related to the proximity of the syllable to the boundary (i.e. syllables closer to the boundary receive most lengthening, and the effect fades out further away from the boundary, as is the case with π -gestures in Byrd & Saltzman 2003), other analyses have suggested that the effect refers to linguistic information. For example, it might be the final-syllable rime that is lengthened irrespective of the number of segments involved. Turk & Shattuck-Hufnagel (2007) investigated the effect of pre-boundary lengthening on American English, and its correlation with pitch accent and stress, and reported that, for American English, lengthening mainly occurs on the pre-boundary rime, but it is also apparent on the whole preboundary syllable, as well as on the rime of the stressed syllable of the preboundary word, thus skipping syllables in between (relevant findings also in Turk & Dimitrova 2007).

The issue of the types of effects exercised by boundary strength on pre-boundary lengthening is, therefore, still an interesting topic that needs data from cross-linguistic literature.

3.2 Goals

The present chapter aims at addressing the following three goals:

1. to provide a methodological tool for confirming that boundary conditions BC0 to BC5 created for this thesis give rise to gradually greater boundary strengths
2. to test how far the phenomenon spreads before the prosodic boundary, and add to the cross-linguistic literature
3. to provide data which will be contextualized with results from the following chapters, especially those from F0 scaling.

3.3 Methodology

3.3.1 Testing for rate effects

In order to control for effects of speech rate on the produced materials, the utterances were given in randomized lists to the participants (different randomization for each speaker). After the production experiment was finished, the produced utterances were tested for any speech rate effects. Speech rate effects mean that a speaker might speak faster at the end of the session than at the beginning, and that change in speech rate might affect the results. In order to test that this wasn't the case, a word was chosen (the word $\chi\rho\acute{\eta}\sigma\eta$ 'use') from the produced utterances that appeared in the same position (close to the beginning of the utterances) in three boundary conditions from the stop voicing dataset. Five One-way ANOVAs were performed, one for each speaker, testing whether the duration of the word was different across conditions. If a significant effect was found, it would suggest that any significant findings found on pre-boundary durations are not due to boundary strength effects, but to speech rate effects.

No effects ($F < 1$) were found in any of the speakers, suggesting that there were no speech rate effects, and that any significant findings are due to boundary strength effects.

3.3.2 Materials

3.3.2.1 Conditions BC1-BC5

For measuring pre-boundary durations the most appropriate segments to use are voiceless stop consonants flanked by vowels, since they offer clear landmarks for segmentation. The template used for creating materials was the one presented in the General Experimental Design chapter of this thesis; the pre-boundary word (always an Adjective) ends in a CV syllable, and the post-boundary word (always a Noun) starts with an open CV syllable, where C is always a voiceless stop (see example (3.1)).

(3.1) [ci'notopi ko'pela] 'common girl' ADJ-NOUN, FEM-ACC SING

1	[a'siðoto pa'zari] 'lawless bargain'	2	[a'siðota pa'zarja] 'lawless bargains'
3	[a'nicito pi'stirio] 'irrefutable proof'	4	[aðj'akopa ka'lezmata] 'endless shouting'
5	[pseftiko pa'gari] 'fake poor-box'	6	[a'fisika ta'ksiðja] 'unnatural travels'
7	[ci'notopi ko'pela] 'common girl'	8	[po'liploci ka'tastasi] 'complicated situation'

Table 3.1: The eight items constructed for investigating pre- and post-boundary durations. Each of those was later embedded in five different prosodic conditions. Neuter nouns in Greek have the same morphological ending in the nominative and accusative plural. The same is true for many feminine nouns in the singular. The nouns used in these experiment had this feature, which means that they could be embedded in prepositional phrases using their accusative forms.

The stress patterns of the pre- and post-boundary words were kept constant, and the sentences were constructed so that the items under investigation did not bear the nuclear accent.

Although the pre- and post-boundary syllables are of main interest, since they are in the vicinity of the boundary, the penultimate and antepenultimate syllables of the pre-boundary word were also designed to be open CV syllables as often as possible, to test how far lengthening spreads in Modern Greek. Eight carrier words were constructed (Table 3.1), and these items were later embedded in sentences for each of the designed boundary conditions BC0 to BC5, as described in Chapter 2.

3.3.2.2 Condition BC0

As mentioned in the General Experimental Design chapter, a set of sentences was also constructed testing a completely different syntactic configuration; that of a definite article in pre-boundary position followed by a noun. The same nouns as the ones used in BC1-BC5, were reused for BC0 (see Table 3.2), and the preboundary article ends in the same vowel, as the Adjective did. For example, for the item [a'jirtiko pa'gari] the corresponding item for BC0 is [to pa'gari]. The two feminine items were excluded for BC0, since their articles in singular accusative end in a nasal /n/, not in an /i/ (e.g. for the item [ki'notopi ko'pela] the resulting item

1	[to pa'zari] 'the bargain'	2	[ta pa'zarja] 'the bargains'
3	[to pi'stirio] 'the proof'	4	[ta ka'lezmata] 'the shoutings'
5	[to pa'gari] 'the poor-box'	6	[ta ta'ksiðja] 'the travels'

Table 3.2: The six items constructed for investigating pre- and post-boundary durations in BC0. Each of those is based on the items constructed for conditions BC1 to BC5.

for BC0 would be [tin ko'pela] 'the girl'), creating non-comparable items, and also potentially giving rise to stop voicing between the pre-boundary nasal and post-boundary stop. Thus, six items were constructed for BC0.

Given that the onset of the preboundary syllable (i.e. the definite article) is always a /t/ (all definite articles in Modern Greek start with an alveolar voiceless stop), no direct comparisons can be drawn between the onset and whole syllable measurements of BC0 and BC1-BC5. However, given that the vowel of the preboundary article was kept the same as the preboundary vowel from BC1-BC5, comparisons of the rime will be reported here. Furthermore, only two male speakers (MS and CP) were still in the recording process when BC0 items for pre-boundary durations and vowel hiatus were created, therefore all results regarding BC0 in this chapter (and in the vowel hiatus resolution chapter) report solely on findings from those two speakers.

3.3.3 Segmentation

The waveform and spectrogram of each file were simultaneously examined on screen for the segmentation of all sounds. The closure and release of the consonant were marked, as can be seen in Figure 3.1. Voiceless stops in Modern Greek are unaspirated, and voiced stops have a negative VOT. The beginning of the closure was marked when there was a clear cessation of F2, and also maybe a clear change in the waveform. The end of the closure is taken as the point where the lips start parting; this is where the measurement of the release starts. The end of the release is located at the point where voicing starts. Figure 3.1 shows a screenshot of

an example of how segmentation was performed. A total of 424 utterances were recorded for this experiment (see Chapter 2, Table 2.1), out of which 31 instances were discarded due to mispronunciations and hesitations. Mispronunciations in any part of the utterance - not just the during the words under investigation - were judged as reasons for discarding an item.

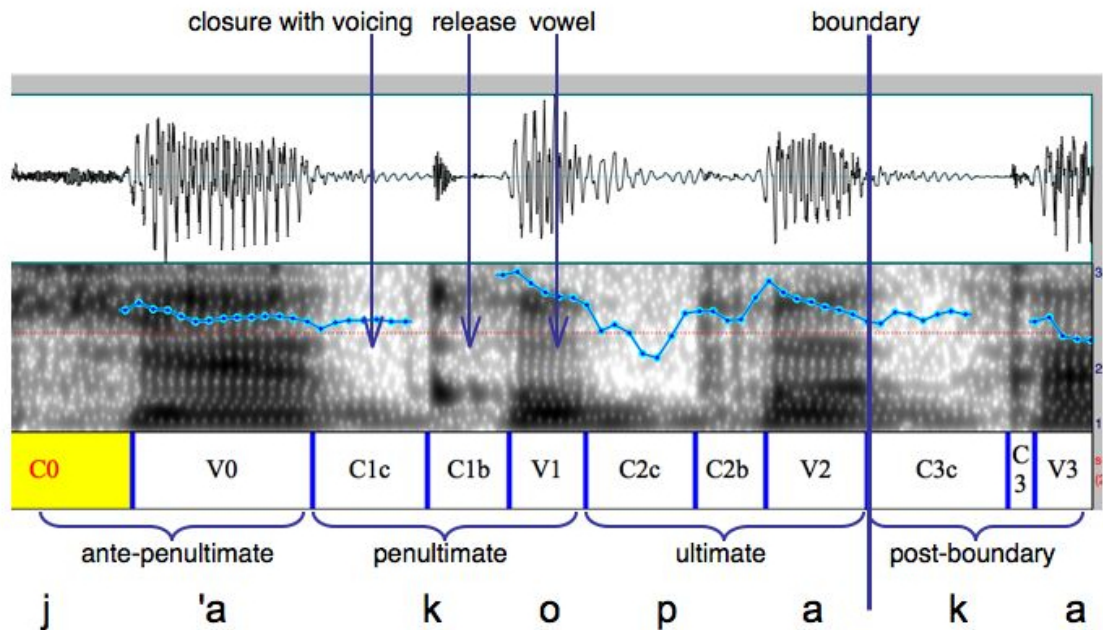


Figure 3.1: Example of segmentation of items constructed for the investigation of pre-boundary lengthening. Sample taken from a female speaker saying the item [a'ðjakopa ka'lezmata]. Figure shows the names and positions of the syllables, and examples of segmentations. It also points to the position of the boundary. This example is taken from BC4.

3.3.4 Measurements

PRAAT (Boersma 2001) was used for segmenting and extracting automatically the measurements of interest (this is the software used for all analyses and segmentations). The durational measurements were extracted automatically from the tier containing the segmentation markers by means of a script. The measurements taken are:

- the duration of the ultimate, penultimate and antepenultimate syllables of the preboundary word, with separate measurements of their onsets and rimes. We

refer to measurements of onsets and rimes instead of segments in the particular instance because the segmental make-up of the onset of the penultimate and antepenultimate syllables of the preboundary word on some occasions comprised of two segments instead of one (see Table 3.1). Therefore, we measure the duration of the whole onset instead of the duration of the individual segments. This does not influence the results, since comparisons are made using repeated measures across conditions.

- the duration of the postboundary syllable, again also separately marking their onset and rimes (the terminology remains onset and rime although all postboundary syllable onsets are made up with one segment, but for continuity reasons with the previous measurements, we still refer to onsets and rimes)
- the duration of lengthening expressed in percent. This means that for each syllable we measured how much each condition was lengthened with respect to BC1. Taking BC1 as a baseline, the amount of lengthening in BC2, BC3, BC4 and BC5 was computed using the following equation:

$$BC_{2345}lengthening = [(BC_{2345}duration - BC1duration) * 100] / BC1duration \quad (3.2)$$

This was done for all syllables, and the results are used to compare the magnitude of lengthening within syllable across conditions, and also across syllables to test how far the phenomenon spreads and how it is distributed.

3.3.5 Statistics

The effect of boundary strength is measured on each dependent variable using 9 REPEATED MEASURES ANOVAS (3 syllables measuring the effect on their onset, rime and whole syllable). Boundary condition with five levels (BC1-BC5) forms the independent variable. The same analysis is also run on the rime of the preboundary syllable with the independent variable having 6 levels (BC0-BC5) for the two speakers that had produced utterances in BC0 (recall that only the preboundary vowel was comparable across all conditions - BC0 to BC5). Differences are regarded to be significant at $p < .05$, and highly significant when $p < .001$, and all post-hoc analyses refer to paired-sample analyses with the Bonferroni correction applied. Three more REPEATED MEASURES ANOVAS are performed testing the

effect boundary strength (four levels, since it is relative to BC1) exercises on the percent lengthening. Finally, three REPEATED MEASURES ANOVAS are used to test boundary strength effect on the duration of post-boundary onset, rime and whole syllable.

3.4 Results

3.4.1 Pre-boundary durations

3.4.1.1 Conditions BC1-BC5

The effect of boundary strength is measured on the onset, rime and whole syllable duration of the ultimate, penultimate and antepenultimate syllables of the final word before the boundary. This tests whether there is a boundary effect and how far it spreads. Table 3.3 shows mean durations for all syllables (onsets and rimes) in each boundary condition, whether the effect was significant, and the results of the Bonferroni paired-samples tests.

The effect of boundary strength was significant for the onset, rime and whole pre-boundary syllable (onset: $F(4, 244) = 37.4$, $p < .001$, rime: $F(4, 244) = 43.4$, $p < .001$, syllable: $F(4, 228) = 53.6$, $p < .001$). The duration of pre-boundary syllables increased gradually across boundary conditions, with BC1 having the shortest duration and BC5 the longest. This supports the claim that the boundary conditions created for this thesis do indeed give rise to stronger prosodic boundaries. Post-hoc results (Table 3.3 last column showing statistically significant differences between conditions) show that all syllables are statistically similar to the condition adjacent to them, but shorter than the one following that; for example, BC1 is shorter than BC2 but not significantly so, while it is significantly shorter than BC3, and similarly BC2 is shorter than BC3 and BC4, but only the difference with BC4 is statistically significant. This gradient variation across boundary conditions shows that the boundary conditions do not give rise to ‘abrupt’ differences; rather, a gradation of possible boundary strengths has been captured by the experimental design. Results are shown in Figure 3.2; each panel shows one syllable, the x axis shows each boundary condition and the y axis the duration of the syllable (ms). It is clear that the ultimate syllable has the most pronounced effect.

BOUNDARY EFFECT ON PRE- AND POST-BOUNDARY SYLLABLES							
	BC1	BC2	BC3	BC4	BC5	Significant	Bonferroni Post-Hoc
ULTIMATE							
onset	67	70	73	80	91	**	123<45
rime	40	45	49	59	76	**	1<3,12<45, 34<5
syllable	105	115	121	139	166	**	1<23<4<5
Percent Lengthening	-	10.5%	17.6%	33.4%	59.9%	**	23<4<5
PENULTIMATE							
onset	65	69	72	73	78	**	1<345,23<5
rime	31	35	35	38	36	**	1<4
syllable	96	103	108	111	115	**	1<2<3, 3<5
Percent Lengthening	-	8.3%	13.6%	16.7%	22.3%	* _(.001)	
ANTEPENULTIMATE							
onset	72	78	83	87	84	-	1<345
rime	69	80	80	84	97	**	1<2345,234<5
syllable	141	158	163	170	181	**	1<2345,23<5
Percent Lengthening	-	14.8%	17.4%	23%	31%	**	
POSTBOUNDARY							
onset	70	73	77	75	68	* _(.012)	1<3
rime	58	55	54	55	57	-	-
syllable	127	128	131	131	127	-	-

*= $p < .05$ **= $p < .001$

Table 3.3: Mean duration (ms) of the onset, rime and whole syllable lengths of the ultimate, penultimate, antepenultimate and post-boundary syllables, along with the percent lengthening of BC2, BC3, BC4 and BC5 compared to BC1 on all pre-boundary syllables.

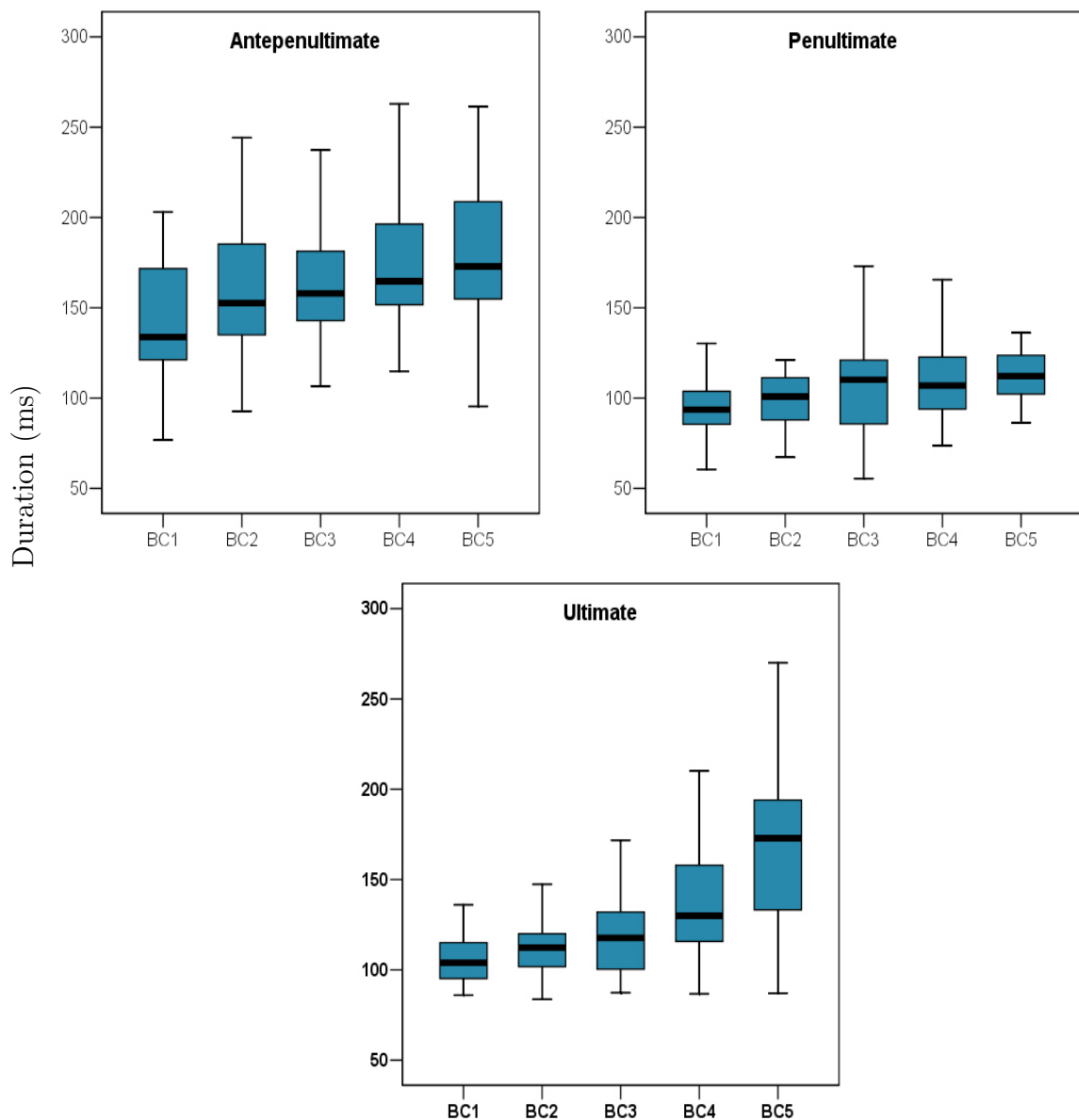


Figure 3.2: Boxplots showing the effect of prosodic boundary strength on duration (ms) of the antepenultimate, penultimate and ultimate syllables of the preboundary word.

A secondary question of this chapter was to test how far the effect of boundary strength spreads before the prosodic boundary. For this reason, the effect was also measured on the duration of the penultimate and antepenultimate syllables. Boundary strength had a significant effect on both the raw durational values and on the percent lengthening of both the penultimate and the antepenultimate syllables (PENULTIMATE onset: $F(4,160) = 10.7$, $p < .001$, rime: $F(4,132) = 4.7$, $p < .001$, syllable: $F(4,132) = 14.3$, $p < .001$, ANTEPENULTIMATE onset: $F(4,128) =$

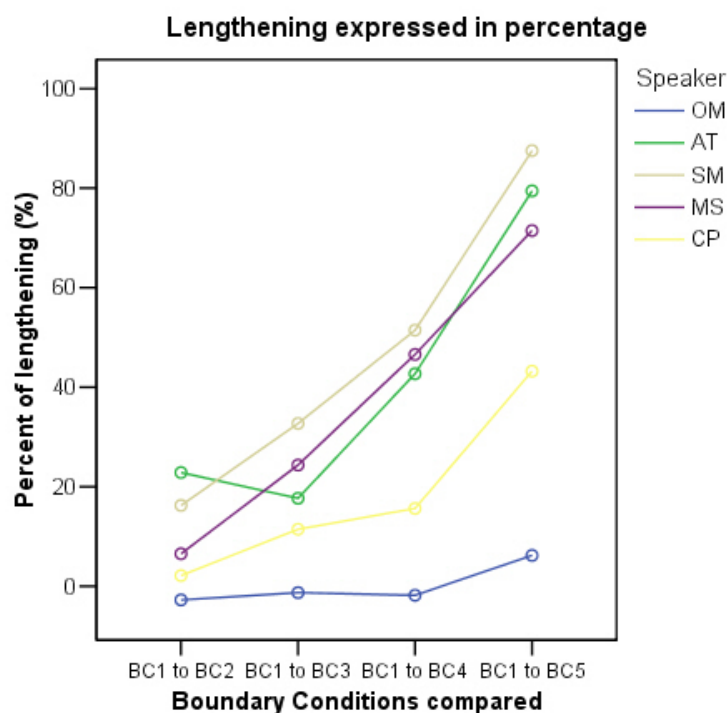


Figure 3.3: Representation of how much the duration of the ultimate syllable was lengthened for each condition

in comparison to BC1 broken down by speaker.

6.38, $p < .001$, rime: $F(4, 128) = 19.55$, $p < .001$, syllable: $F(4, 128) = 22.2$, $p < .001$, PERCENT LENGTHENING PENULTIMATE SYLLABLE $F(3, 99) = 5.97$, $p = .001$, PERCENT LENGTHENING ANTEPENULTIMATE SYLLABLE $F(3, 96) = 9.5$, $p < .001$). The effect of boundary strength was, therefore, apparent in all syllables reaching up to the antepenultimate. Given the present design it is not possible to test before this syllable, since the segmental make-up was not controlled for. The effect on the penultimate and antepenultimate syllables is similar to that on the ultimate syllable, i.e. the duration of the syllables is getting longer higher in the structure (see Figure 3.2).

Having identified that there is lengthening in Modern Greek, and that the effect reaches up to the antepenultimate syllable at least, we asked the question of whether there is a difference in magnitude in how pronounced the effect is across syllables.

This was tested by investigating the percent lengthening in each syllable. We tested the amount of lengthening that took place from BC1 to BC2, BC3, BC4 and BC5. Figure 3.3 presents the amount of lengthening on the pre-boundary syllable of each condition in relation to BC1 for each speaker. BC2 was lengthened by 10.5% in comparison to BC1, BC3 was lengthened by 17.6%, BC4 by 33.4%, and finally BC5 by 59.9%. Boundary strength effect on the percent lengthening of the pre-boundary syllable was significant (ULTIMATE $F(3, 171) = 43.191, p < .001$), and the Bonferroni paired samples t-tests showed that the differences across BC2/BC3 vs BC4 vs BC5 were statistically significant showing that the experimental manipulations for each of those boundary conditions are indeed prominent, and can be reliably used for the rest of the thesis as means of identifying boundary strength effects.

Importantly, the effect is most pronounced on the ultimate syllable; the percent lengthening starts from approximately 10% and reaches up to 60%, while for the penultimate syllable the percent of lengthening starts from approximately 8% and only reaches up to 22%, and in the antepenultimate, the percent lengthening starts from 15% and reaches up to 31%. Moreover, it is also clear that the effect is more pronounced on the *antepenultimate* than the penultimate syllable, that is the syllable further away from the boundary. This signals a large difference between the two syllables. Hence, the effect is most pronounced on the pre-boundary syllable, followed by the antepenultimate and then by the penultimate syllable. Given that the antepenultimate syllable is the one carrying the stress on the preboundary word, it suggests that the effect is more attracted to the stressed (antepenultimate) syllable, however it definitely does not completely skip the intervening unstressed syllable (as was found for English in Turk & Shattuck-Hufnagel 2007)².

3.4.1.2 By speaker analysis

The effect of boundary strength on the duration of the preboundary syllable was also measured for each speaker separately using a Univariate ANOVA. The effect of speaker was found to be significant ($F(4,336)=14.693, p < .001$) and there was a significant interaction between boundary conditions and speaker ($F(16,336)=3.752,$

²Although other things could be attributing to the effect, it is clear that there is a difference between the ante- and the penultimate syllables, supporting Turk & Shattuck-Hufnagel's (2007) finding.

$p < .001$), effects which stemmed from the fact that speaker OM did not show any lengthening patterns (see Figure 3.3). This suggests that lengthening is not a cue that is obligatorily used by speakers to mark boundary strength, while the majority of speakers make use of it. During the course of this thesis, results of by speaker analyses will be reported to test how speaker OM marked boundary strength since she did not use pre-boundary lengthening.

Separate One-way ANOVAs were performed for each speaker testing for differences on how they grouped boundary conditions. Figure 3.4 shows the results of Bonferroni post-hoc tests on how each speaker grouped statistically the five boundary conditions. The lines across boundary conditions show the ones that were not statistically different, while when two adjacent conditions are not part of the same line, it means that the two conditions were statistically different. For example, speaker AT grouped conditions BC1, BC2 and BC3 on one hand, and conditions BC2, BC3 and BC4 on the other. This means that conditions BC1 and BC4 were statistically different, and it is represented in the graph by the fact that BC1 and BC4 do not co-occur in the same line. BC5 for that speaker was statistically different to all previous ones, and it is marked by a line on its own.

The general picture taken from the types of groupings performed by each speaker is that there is gradual lengthening from lower to higher conditions. BC1 and BC2 are always statistically distinct from BC4 and BC5 (with the exception of speaker AT for whom BC2 and BC4 were not statistically distinct). BC3 is always in the middle between BC2 and BC4, sometimes grouping with BC2, sometimes with BC4, and always being different to BC5. Therefore, there is a clear difference between lower and higher domains, with BC1 and BC2 being distinct from BC4 and BC5, but there are also intermediate differences, which are very interesting too. The difference between conditions BC2 and BC3 is a marked syntactic manipulation, whereby the intervening boundary is one between a prepositional phrase and a verb phrase in BC2 vs a clause boundary in BC3 (see Chapter 2 for the construction of boundary conditions), which however is not always signalled durationally. Similarly, the difference between BC3 and BC4 is a marked difference of strength, which sometimes is and sometimes isn't marked durationally. However, both effects can be marked, and the effect across types of manipulations (syntactic vs length) gives

rise to the same type of effect, i.e. a lengthening effect. That is, two completely different manipulations can give rise to similar effects on duration.

These results show that pre-boundary lengthening does mirror boundary strength, but they also suggest that there are differences across speakers. This finding can either be due to different speakers grouping prosodic domains differently (having a different structure representation), or it can more easily be attributed to the fact that pre-boundary lengthening is not meant to provide clear support for the existence of one labelled domain over the other, but rather is solely a clear correlate of boundary strength, as it was taken to be by the direct syntax mapping approach. However, the length manipulation shows that the direct syntax mapping approach can't capture all phenomena, supporting the necessity of postulating an intermediate prosodic structure.

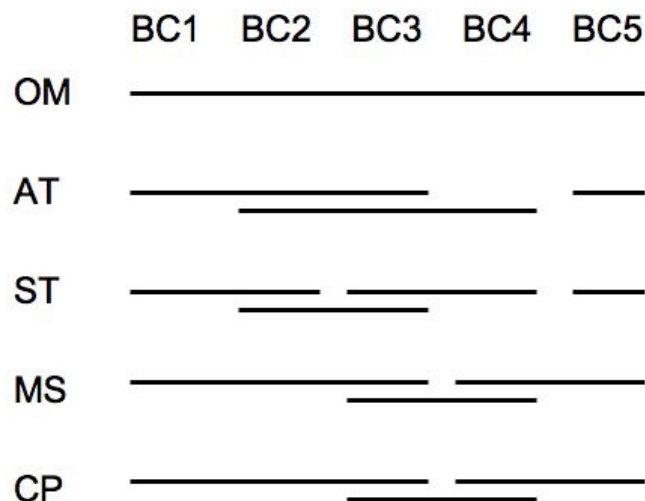


Figure 3.4: Schematic representation of how each speaker signalled the effect of boundary condition. Domains that stretch within the same line are statistically similar, while domains belonging to different lines are statistically distinct.

3.4.1.3 Conditions BC0-BC5

As mentioned in the General Experimental Design chapter, after some of the recordings took place, extra materials were added representing a lower domain than BC1, namely BC0, mainly for the investigation of post-nasal stop voicing, but for comparison purposes materials were created for all phenomena. Only two speakers were recorded producing the materials for this part; MS and CP - both males.

BOUNDARY EFFECT ON PRE-BOUNDARY RIME INCLUDING BC0								
	BC0	BC1	BC2	BC3	BC4	BC5	Significant	Bonferroni
ULTIMATE								
rime	46	45	52	58	69	73	**	012<45

*=p<.05
**=p<.001

Table 3.4: Mean duration (ms) of the pre-boundary rime for each of the six prosodic conditions including BC0 for only the two male speakers that were recorded producing the items.

Repeated measures analyses on the duration of the preboundary rime³ showed a significant boundary strength effect ($F(5, 45) = 10.144, p < .001$). It was found that BC0 grouped with BC1, BC2 and BC3, being statistically different to BC4 and BC5. A univariate ANOVA testing the interaction between boundary conditions and speaker revealed that there was no interaction, and Bonferroni post-hocs from one-way ANOVAs on each speaker separately showed that the two speakers grouped all conditions in the same manner, with BC0 not being shorter or longer than BC1 for either one of them. Table 3.4 reports on the mean values of the rime of the pre-boundary syllable for conditions BC0-BC5, and Figure 3.5 shows the lengthening effect of prosodic boundary strength for those two speakers, including BC0.

The difference (syntactic and prosodic) between BC0 and BC1 is quite marked, and it is interesting that it is not distinguished durationally by the two speakers. This result will be compared to results from other phenomena in this thesis.

3.4.2 Post-boundary durations

Previous research has shown that the duration of post-boundary segments is also influenced by the strength of the boundary, although the exact effect depends on the language and the segments investigated (e.g. Cho & McQueen 2005, Kuzla 2009). For the materials investigated in this experiment, the effect of boundary strength was found to be significant for the onset of the postboundary syllable ($F(3, 180) =$

³Remember that in the pre-boundary syllable the only possible comparison between BC0 and the other conditions is the duration of the rime, since the onset of the preboundary syllable was always a /t/.

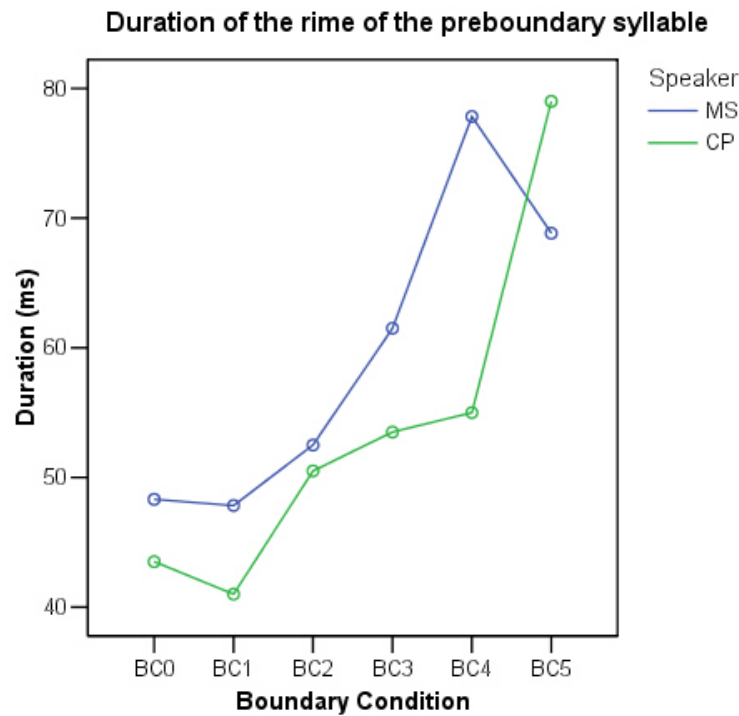


Figure 3.5: Mean duration of pre-boundary rime (*i.e.* vowel) by speakers MS and CP for renditions BC0-BC5 ($n=144$).

3.76, $p=.012$), and non-significant for the rime and whole syllable. Paired-sample tests with the Bonferroni criterion applied on the duration of the onset showed that the only significant difference in all comparisons was between BC1 and BC3. This suggests that the effect is not gradually increasing, and the differences in means across conditions (as shown in Table 3.2) are not different.

It seems that Modern Greek does not exhibit post-boundary lengthening for phrasal categories, at least as far as the durations of the post-boundary segments is concerned. This does not exclude the possibility that initial-strengthening is used in some other articulatory fashion, but this has not been the matter of the present investigation. However, from the present results it is clear that the pre- and post-boundary segmental production in terms of duration patterns is influenced differently by boundary strength, with the post-boundary segments not showing any

effects, while the pre-boundary syllables are affected up to the antepenultimate stressed syllable.

3.5 Conclusions

The goal this chapter set out to achieve was to test whether the boundary conditions created for measuring boundary strength effects have managed to capture differences of boundary strength starting from weaker and moving to stronger across BC0 to BC5. It was hypothesized that, if there is a boundary strength effect, then BC0 will be the weakest boundary showing the shortest pre-boundary durations, and BC5 would be the strongest, showing the longest duration. In addition, it was hypothesized that intermediate conditions would have variable durations, since it is anticipated that pre-boundary durations get gradually longer higher in the structure, and that is there isn't a binary distribution separating the lowest from the highest conditions.

These hypotheses were confirmed by the data; longer durations were found in higher boundary conditions, and the distribution of values across conditions changed gradually, in the sense that there was no abrupt change from one condition to the other, nor a binomial distribution separating one group of conditions with respect to another group. Therefore, it is confirmed in this chapter that the experimental design has succeeded in producing varying boundary strength conditions, which get gradually stronger. Moreover, lengthening was found to be an optional technique for marking boundary strength, since one out of five speakers did not lengthen pre-boundary segments at all.

A secondary goal that this chapter set out to investigate was the question of how far the effect of boundary strength spreads before the prosodic boundary; one view is that its effect is most prominent close to the boundary, and fades out the further away (like the proposal put forward by the π -gestures within the Articulatory Phonology theory Byrd & Saltzman 2003), while a different one proposes that it is a linguistic variable that gets affected, like the pre-boundary rime or pre-boundary syllable (Wightman et al. 1992), while a final view is that both types of

events can co-occur depending on the expandability of the final syllable (Cambier-Langeveld 1997)⁴. Since the experiment was not designed to answer this question, only partial answers can be provided, and the basis for further experimentation in Modern Greek.

Results showed that the effect of boundary strength was significant on all three pre-boundary syllables measured (ultimate, penultimate and antepenultimate), but there were differences in magnitude on how much lengthening took place. The effect appeared to be the strongest on the ultimate syllable, with BC5 lengthening by 59% in comparison to BC1. Subsequently, the effect was most significant on the *antepenultimate* syllable, which was also the stressed syllable. BC5 showed a 31% lengthening in comparison to BC1 for the antepenultimate syllable, while the intervening unstressed penultimate syllable lengthened its BC5 by 22% with respect to BC1. This finding supports a mixed effect of boundary strength on the syllable immediately preceding the boundary, but also on linguistic variables like the accented syllable (supporting findings from Turk & Shattuck-Hufnagel 2007 and Cambier-Langeveld 1997). Modern Greek does not have a distinction between long and short vowels, and the durations of its vowels more closely resemble short English vowels (or have duration values inbetween long and short English vowels but more closely resembling short vowels Lengeris 2009), therefore it is a matter of further investigation to see whether the fact that the vowels might be ‘expandable’ in this language is one factor allowing for extra linguistic variables to be lengthened before the boundary.

In addition to the difference in magnitude across syllables, there was also a difference in the number of significant distinctions each syllable participated in across the five boundary conditions. Post-hoc analyses on the pre-boundary syllable showed that statistically at least four domains were distinct (BC1<BC2 and BC3<BC4<BC5), while in the antepenultimate syllable there was only one statistically significant difference (BC1 being the shortest), with gradient intermediate values for the rest. This suggests that, while the effect is larger on the antepenultimate syllable than the pre-boundary one, there is still some difference across positions. However,

⁴A full discussion and terms proposed for naming each of those proposals has been presented in Turk & Shattuck-Hufnagel (2007).

the present design was not designed to test those questions, therefore it is clear that more experiments are in need to elaborate on the current findings on how boundary strength effects are distributed on the pre-boundary word, and on the factors influencing this distribution. It seems that Modern Greek can offer an interesting setting for investigating these questions.

Interestingly, boundary strength effects seem to be realised differently in pre- vs post-boundary positions in Modern Greek. While there is a boundary strength effect that spreads until at least the antepenultimate syllable, in post-boundary positions the effect is barely visible on the onset of the post-boundary syllable. This goes against an explanation, as offered by the current version of the π -gestures theory, which states that the gesture is centered on the boundary and spreads on both sides of it getting gradually weaker (Byrd & Saltzman 2003); under this view, the boundary effect should spread in post-boundary positions in the same manner, as it does in pre-boundary positions, but that is not the result found in this experiment.

The current findings prove that Modern Greek does exhibit final lengthening above the word level, and seems to not have any lengthening (at least on the two speakers having produced both BC0 and BC1 items) across prosodic word junctures, supporting findings from previous research (Arvaniti 2000, Katsika 2007). Moreover, they suggest that Modern Greek does not have initial strengthening realised in terms of durational lengthening as prominently as other languages have, at least on voiceless plosives; however, more investigation is in order.

One final issue has to do with whether pre-boundary durations provide support for distinct types of labelled domains in the prosodic structure. The differences across speakers in the types of groupings formed, as well as the great variation in each boundary condition, make it difficult for a researcher to propose any specific *types* of domains on the basis of durational patterns. In the introduction to this chapter it was discussed that pre-boundary durations have mainly been used in tandem with other correlates to the structure in postulating specific domains within the structure. This chapter has clearly shown the large amount of variability in marking boundary strength by means of duration, highlighting that, while boundary

strength effects can easily be identified, it is not possible to distinguish whether they are derived from distinct types of prosodic constituents that are just gradually produced on their acoustic manifestation, or whether they are the result of ranging boundary strengths of indeterminate depth. For this reason, the present chapter has only been used as a methodological tool for confirming the experimental design, and the following chapters turn to the exact manifestation of boundary strength on correlates that are traditionally believed to be able to distinguish categorically across different types of domains. Moreover, this chapter provides data against which findings from F0 scaling will be compared in Chapter 6. Finally, it is shown that both syntactic and length manipulations can give rise to changes in boundary strength (e.g. differences between BC2 and BC3 vs BC3 and BC4), a finding that will be further discussed in the final chapter of this thesis, Chapter 7.

CHAPTER 4

Post-nasal Stop Voicing

Sandhi processes have been viewed as a major correlate of prosodic structure, especially within labelled views of prosodic constituency. The phonological nature of several sandhi processes, however, has been questioned by recent research showing that many processes traditionally treated as sandhi are actually the result of coarticulation. Additionally, labelled views using sandhi-type phenomena for supporting the existence of specific domains have relied on the assumption that boundary strength effects will be qualitative, either allowing or blocking the application of the process across prosodic boundaries.

Separating sandhi (or sandhi-type) phenomena from coarticulatory processes when identifying prosodic structure effects on their application is important, because the two might be influenced differently by boundary strength, and there can be different methodologies for measuring boundary effects across processes. For example, the effect of boundary strength on sandhi processes can be defined by how many times the phenomenon occurs within each boundary condition, while in the case of coarticulatory processes by the amount of coarticulation taking place within each condition. Moreover, it has been found that boundary strength effects are not always mirroring qualitative boundary strength effects, and can be variable on how many times a sandhi process takes place across boundary conditions.

It is a matter of research, therefore, to see if there are sandhi processes that can provide solid evidence for labelled domains using acoustic measurements.

This and the following chapter focus on two sandhi-type processes, post-nasal stop voicing and vowel hiatus resolution, with the goal of first testing whether their acoustic manifestation takes place in a categorical or gradient fashion (indirectly supporting whether their nature is phonological or coarticulatory), and then establishing the nature of potential boundary strength effects on them.

There are several reasons motivating the analysis of post-nasal stop voicing. First, it has been used in past research within labelled views (Nespor & Vogel 1986) to support the existence of labelled prosodic domains. Second, Nespor & Vogel's (1986) analysis of the process (like all analyses of the phenomenon thus far) assumed it to be a sandhi phenomenon (not the result of coarticulation), when acoustically this has not been tested. Third, Nespor & Vogel's (1986) analysis of the effect of boundary strength on the phenomenon has relied on impressionistic data relating to a very limited number of prosodic conditions; the present chapter aims at investigating the phenomenon using a perception and production experiment in a variety of boundary conditions. The perception experiment is meant to test previous impressionistic analyses by looking for boundary strength effects on how many times raters classify the phenomenon taking place across boundary conditions, and if it seems to undergo a qualitative boundary strength effect on the basis of such ratings. The production experiment is meant to (i) record voiceless and voiced stops to identify variables that are significant for the voicing distinction, therefore establishing what a voiced stop is in Modern Greek, and (ii) test boundary strength effects on 'nasal+stop' sequences for qualitative effects of boundary strength that would support labelled views.

Finally, the application of this phenomenon will be compared to that of vowel hiatus resolution, comparing boundary strength effects on a phenomenon thus far viewed as phonological (i.e. stop voicing) with one found to be coarticulatory (i.e. vowel hiatus resolution).

4.1 Introduction

4.1.1 Definition of Post-Nasal Stop Voicing in Modern Greek

Post-nasal stop voicing is the process whereby a voiceless stop (/p,t,k/) is replaced by its voiced counterpart (/b,d,g/ respectively) when following a nasal consonant (/m,n/). It is a phenomenon that has been attested in several languages, like Modern Greek (Newton 1972), Japanese (Ito & Mester 1986) and others (see Hayes & Stivers 2000, for references on more languages).

In Modern Greek, clusters giving rise to stop voicing can be found both lexically and post-lexically. *Lexically*, they can occur word-initially and word-medially; for example, word-initially ντάμα ‘checkers’ (which can be pronounced either as [ˈdama] or [ˈⁿdama]), and word-medially σόμπα ‘stove’ (which can also be pronounced with the pre-nasalised and non-pre-nasalised rendition). When such clusters are found within a lexical item, voicing is obligatory (but the pre-nasalisation is optional, and has been found to be influenced by social factors and style, e.g. Arvaniti & Joseph 2004), as in the following example:

(4.1) έντιμος → [ˈedimos] ‘honourable’

Post-lexically, voicing the stop is optional. The regulating factor determining the application of stop voicing has been thought to be the prosodic structure. When the ‘nasal+voiceless stop’ sequence is separated by a word boundary, it has been suggested that the speaker has three options (Arvaniti 1999); to keep the nasal and not voice the voiceless sound, keep the nasal and voice the voiceless stop, or delete the nasal and voice the voiceless stop (example 4.2).

(4.2) τον πατέρα ‘the father’ MASC-SING-ACCUS
 no voicing: [ton paˈtera]
 voicing with prenasalisation: [to^mbaˈtera]
 voicing with no prenasalisation: [to^baˈtera]

Therefore, voicing the post-nasal stop is optional, and it is the issue of this chapter to identify the effect of boundary strength on whether it will take place or not.

Before switching to boundary strength effects on the phenomenon, however, we first review the nature of post-nasal stop voicing itself. The above description presented three possible outcomes for ‘nasal+stop’ sequences; using such transcriptions could lead to the conclusion that those are three phonologically distinct processes. However, it is possible that all are results of coarticulation with one end having no voicing, and in the other end having full voicing and nasal deletion, while in between cases, for example, keep the nasal and voice the stop in a gradation of possible coarticulatory renditions, analogous to the analysis of assimilation in English (Nolan et al. 1996). This relates to the present investigation, since it suggests that the acoustic difference between a voiceless stop and a voiceless stop having undergone voicing might range in a continuum of possible values which are phonetic, and do not relate to a phonological distinction. This would suggest that the voicing of the stop is not the result of a phonological sandhi process, but of coarticulation.

The issue of prenasalization is not addressed in the present thesis, since there have been no nasal and oral airflow measurements. Arvaniti & Joseph (1999) reported that instances of stop voicing might be accompanied by deletion of the pre-boundary nasal. They propose that the application of the phenomenon is categorical, i.e. that deletion either takes place or not parallel to voicing, and that the behaviour of the pre-boundary nasal might also act as a cue to voicing distinctions. Although the present experiment is not testing the behaviour of nasality, a preliminary impressionistic analysis of classifying where nasal deletion might have occurred is performed.

Post-nasal stop voicing in Modern Greek is a very interesting process for several reasons. First, it has not been investigated acoustically in the past; all previous analyses of the phenomenon have analysed it impressionistically, a gap that this chapter aims at addressing. These impressionistic analyses have reported that the phenomenon is affected by AGE, in that younger speakers almost always produce the voiced rendition (if the prosodic context is right), while style and education have not been reported as important factors (Arvaniti & Joseph 1999). This is the reason why this thesis only uses speakers in their twenties.

Given that there is no research on the acoustics of post-nasal stop voicing in Modern Greek, the next section discusses differences between voiceless and voiced sounds in Modern Greek. Following that is a presentation of how post-nasal stop voicing in Modern Greek has been used in the past by labelled views to propose the existence of specific labelled domains.

4.1.2 Phonology and Phonetics of voiceless and voiced stops in Modern Greek

Modern Greek has bilabial, alveolar and velar voiceless and voiced stops (Mirambel 1961, Mackridge 1985, Philippaki-Warburton 1992, Arvaniti 2007). Voiceless stops are said to be unaspirated (Fourakis 1986*b*, Botinis, Fourakis & Prinou 1999, Arvaniti 2001*b*, Nicolaidis 2002), and voiced stops have a negative VOT, with full prevoicing during the closure.

While many studies have investigated the phonetics of *voiceless* stops of Modern Greek, mainly focusing on their durational patterns (Fourakis 1986*a*, Nicolaidis 2001, Nicolaidis 2002), no research to our knowledge has looked into the acoustic differences between voiceless and voiced stops. Table 4.1 is adjusted from Arvaniti (2007) and shows the mean durations reported in some studies investigating the duration of the closure and VOT of voiceless stops in Modern Greek¹ (the table has been adjusted to present data from an experiment performed in this thesis testing voiceless and voiced stops acoustic differences, presented in Section 4.3.1). All these studies agree that /k/ has the shortest closure duration, and the longest VOT, while /p/ has the longest closure (see Cho & Ladefoged 1999, for universal tendencies in VOT patterns).

Kong, Beckman & Edwards (2007), one of the very few studies investigating some aspect of voiced stops in Modern Greek, looked at the acquisition of voiced stops by children, and reported that durational and amplitude differences can distinguish between voiceless and voiced stops of Modern Greek. For this reason amplitude measurements are taken along with duration in the present chapter. However, their

¹Given that they used different materials, speech styles, and experimental conditions, there are differences in the raw values they report.

results were based on the analysis of solely one adult speaker, and it is not known to what extent they generalize to other members of the population.

		[p]	[t]	[k]	[b]	[d]	[g]
Closure	Fourakis (1986 <i>b</i>)	113	101	90			
	Arvaniti (1987)	88	85	86			
	Arvaniti (2001 <i>a</i>)	56	50	44			
	Nicolaidis (2002)	69	63	**			
	Kainada (this thesis)	63	43	45	65	70*	58*
VOT	Fourakis (1986 <i>b</i>)	9	16	23			
	Arvaniti (1987)	11	15	26			
	Arvaniti (2001 <i>a</i>)	13	16	23			
	Nicolaidis (2002)	14	22	**			
	Kainada (this thesis)	18	23	29	11* ^a	13* ^a	16* ^a

* $p < .001$ - the voiced segment significantly different to its voiceless counterpart

** Nikolaidis (2002) did not look at velar stops

^aThese measurements represent the duration of the release, not the actual VOT which is expected to be fully prevoiced

Table 4.1: Table showing the mean values of closure and VOT of all Greek voiceless stops, as presented in previous research, with results from this thesis added. Table adjusted from Arvaniti (2008). Results for voiced stops only come from this study, and what is marked as VOT refers to the duration of the release, and not the voice onset, since voiced stops in Modern Greek are fully prevoiced.

Research on voiceless stops has also shown that, in connected speech, there can be residual voicing from the preceding vowel in the closure of the stop. Nicolaidis (2001) and Nicolaidis (2002) report that the closure of the voiceless /t/ and /k/ in spontaneous speech was often partially or fully voiced in intervocalic position, or between a vowel and voiced consonant. Nicolaidis (2001) showed that there is an amount of carryover voicing from the preceding vowel in the closure of the following stop. This is important to note, since this carryover effect does not appear to influence the phonological status of the voiceless segment (turning it into a voiced segment), but rather seems to be phonetic in nature. Lisker & Abramson (1964) emphasize the universality of the fact that this carryover effect is not expected to affect the phonological status of the following segment. They specifically mention that laryngeal oscillations of preceding voiced environment are possible to be found in the closure of the stop, that these do not influence the phonological status of the consonant, and they should be called “edge vibrations”. In this thesis the amount of carryover voicing from the preceding nasal sound was examined as a

factor potentially influenced by prosodic boundary strength, and also influencing the voicing of the stop, and it was found that while the extent of carryover voicing from the preceding voiced sound was influenced by boundary strength (less carryover voicing in higher order domains), it did not influence the perception of a sound as voiceless or voiced, thus supporting Nicolaidis's (2001) and Lisker & Abramson's (1964) analysis. For this reason these results are not discussed further.

From the above it becomes apparent that more research is needed on the acoustic and articulatory characteristics separating voiceless from voiced stops in Modern Greek, as well as the phonetic characteristics of sounds undergoing voicing after a nasal. A production experiment is performed in this chapter first, in order to identify acoustic differences between voiceless and voiced stops in Modern Greek. This will act as a baseline against which the acoustic characteristics of 'nasal+voiceless stop' sequences can be compared, in order to specify whether voicing has taken place. Both the voiceless and voiced stop and the 'nasal+voiceless stop' datasets are recorded in a variety of boundary strength conditions. This way it is possible to form expectations as to whether findings on 'nasal+stop' sequences are solely due to boundary strength effects like initial strengthening, or due to voicing of the post-boundary sound.

4.1.3 Prosodic structure and post-nasal stop voicing in Modern Greek

It is well established by now that sandhi processes are considered as one of the main correlates in identifying prosodic domains. A sandhi process is viewed as a marker of prosodic constituency when allowed it is across a prosodic boundary and blocked by a stronger boundary, therefore providing binary distinctions between domains. Such an effect is taken to be a qualitative boundary strength effect mirroring distinct domain types within the prosodic structure. Post-nasal stop voicing is investigated here to test whether the segmental change from voiceless to voiced stop is phonetically gradient, and to see how boundary strength affects its application, i.e. whether a qualitative or variable boundary strength effect can be identified.

As noted in the introduction to this chapter, post-nasal stop voicing has been specifically selected in this thesis because it was been one of the processes used in past research by Nespor & Vogel (1986) to support the existence of qualitatively

distinct domains within the prosodic structure. The use of a previously visited phenomenon allows us to test the validity of previous proposals, while concurrently testing the actual application of such processes. Nespor & Vogel's (1986) analysis has been very influential within the realm of prosodic structure research, and many researchers have relied on their analysis of prosodic constituency.

In their analysis post-nasal stop voicing has been viewed as a phonological rule changing a voiceless stop to its voiced counterpart, and it has been suggested that the phonological rule regulating its application relies on the prosodic grouping of the words involved. Therefore, Nespor & Vogel (1986) propose that (i) post-nasal stop voicing is obligatory within a phonological word, (ii) optional within a clitic group across two phonological words, and (iii) blocked across clitic groups.

- a. [-cont] → [+vce] / [... [+nas] __...]_ω (obligatory)
 Obligatory application within phonological word
 έντιμος ['edimos] 'honoured'
- b. [-cont] → [+vce] / [... [... [+nas]]_ω [__...]_ω...]_C (optional)
1. Optional application across phonological words within clitic group
 δεν περιάζει [ðembi'razi] (<[ðen][pi'razi]) '(it) doesn't matter'
 2. Blocking across clitic groups
 όταν πάς ['otan pas] (*['otam bas]) 'when (you) go'

Note the difference between examples (4.1.3b. 1 and 2); Nespor and Vogel propose that the negative particle [den] is a phonological word grouping with the following word to form a clitic group, while the temporal particle ['otan] is said to be a clitic group on its own. In Nespor and Vogel's analysis it is not clear what the difference is between these two items, giving rise to the different prosodic groupings. The present chapter specifically addresses this issue by using a variety of particles, with the goal of trying to shed some light on why they propose this distinction, and how other particles might behave.

As with all sandhi processes investigated in their book, Nespor & Vogel's (1986) analysis is faced with the question of whether the processes they have been treating as phonological are indeed phonological or coarticulatory, like assimilation and vowel hiatus resolution have been found to be in other studies (Nolan 1992, Nolan et al.

1996, Zsiga 1997). Additionally, the issue of whether boundary strength effects can give rise to qualitative effects across prosodic contexts is addressed, as the main point of interest.

Since Nespor & Vogel's (1986) is the only study suggesting that post-nasal stop voicing can act as prosodic constituency marker, the present chapter aims at directly testing their proposal. Therefore, first a pilot perception experiment will test the validity of such impressionistic analyses, and the type of boundary effects surfacing from perception experiments on stop voicing. Subsequently, the production experiment will identify the acoustic manifestation of the phenomenon, and the effect boundary strength exercises on the acoustic differences between voiceless and voiced stops.

4.1.4 Prosodic structure influences on stop consonant production

The application of a sandhi process presupposes the change from one sound to another; in the case of post-nasal stop voicing it is presupposed that a voiceless stop in post-boundary position will change to a voiced stop. Before investigating the actual boundary strength effect on the application of the sandhi-type process, it is important to know how voiceless and voiced stops are influenced by boundary strength in the absence of the sandhi site. This allows us to identify the effects boundary strength exercises normally on stop consonants in the absence of a sandhi site, and therefore distinguish boundary effects that stem from, say, initial strengthening, to those stemming from the application of the sandhi phenomenon.

Prosodic structure is well-known to affect the duration of post-boundary segments (Fougeron & Keating 1997, Keating 2003) by means of 'prosodic strengthening' in many languages (see Chapter 3). Much research on the phonetic and articulatory characteristics of boundary strength effects on post-boundary positions has focused on stop consonants. The nature of the post-boundary effect varies across languages, and seems to affect different phonetic and articulatory characteristics of the voiceless and voiced sounds. In English, VOT of stops is influenced hierarchically with longer VOT in higher domains (e.g. Pierrehumbert & Talkin 1992, Choi 2003, Cole, Kim, Choi & Hasegawa-Johnson 2007), and the duration of the *closure* of /t/s is longer in high prosodic conditions Cho & McQueen (2005). The duration of the VOT

of Dutch alveolar voiceless stops, on the other hand, is shorter in higher prosodic positions, showcasing a different manifestation of boundary effect between English and Dutch. Korean alveolar voiceless stops and nasals are influenced hierarchically with lengthening effects throughout the segment (Cho & Keating 2001, measured on a variety of articulatory and acoustic parameters like linguopalatal contact, duration of stop closure and VOT, total voiceless interval). Hayashi et al. (1999) reported a clear prosodic effect on the amount of linguopalatal contact and closure duration of alveolar voiceless unaspirated stops of Taiwanese, but not on its VOT, while Fougeron (1999) reported a prosodic effect of boundary strength on French VOT, but not as clear as the one for the other studies just mentioned. The same stands for Choi (2003), who reported that, for English, the VOT of /p/ and /b/ consonants was not affected by prosodic strengthening as significantly as it was by focus.

With respect to Modern Greek, there has been no research investigating boundary strength effects on pre- and post-boundary durations per se. In Chapter 3 we tested the effect boundary strength exercises on the timing patterns of boundary adjacent segments, and found that Modern Greek does exhibit final lengthening, while in post-boundary positions the effect is very small and generally only reaches the very first post-boundary segment. Given that the post-boundary segments in that experiment were voiceless stops, they are similar to some of the items in the present experiment. The anticipated boundary strength effect, therefore, is expected to be small on the post-boundary stop, and in particular for voiceless stops it is anticipated to give slightly longer durations higher in the structure.

4.1.5 Research Goals and Chapter outline

The goals of the present study are:

- to test whether the acoustic manifestation of post-nasal stop voicing takes place categorically or gradiently, indicating a sandhi or a coarticulatory phenomenon
- to test the effect of boundary strength on the application of stop voicing with the goal of identifying whether it offers support for proposing that the prosodic structure entails labelled domains. This will be accomplished using

three techniques: (i) by first identifying the acoustic differences across voiceless and voiced stops, creating expectations as to how ‘nasal+stop’ sequences having undergone voicing should look, (ii) by testing boundary strength effects on voiceless and voiced stops so as to form expectations as to how boundary strength effects look in the absence of a sandhi site, and (iii) by finally investigating the acoustic manifestation of ‘nasal+voiceless stop’ sequences, while knowing what voiced stops look like, and what the acoustic manifestation of the variables measured should look like if there is a qualitative boundary strength effect giving rise to voicing instances in some conditions and not others

- by answering the previous goal, we will also be testing whether Nespor & Vogel’s (1986) analysis of post-nasal stop voicing application was correct. This will be accomplished by elaborating on the items used in the lowest boundary condition in comparison to those used by Nespor and Vogel
- to provide the first phonetic analysis of stop voicing in Modern Greek, while on the side providing data for comparison with those in the next chapter, where we investigate vowel hiatus resolution.

The goals of the present experiment are, therefore, addressed using three experiments:

1. **A preliminary perception experiment**

This is the experiment testing whether perceptually the impressionistic analyses of Nespor and Vogel stand. Moreover, it will separate the instances that are perceptually classified as voiceless and those classified as voiced, which will be analysed phonetically to see if they are categorically distinct, or varying gradiently. If they are distinct, it will provide initial support that the phenomenon is indeed a sandhi process.

2. **A production experiment on acoustic differences between voiceless and voiced stops** and the effect boundary strength exercises on them.

This experiment will serve two purposes. First, it will investigate how voiceless stops differ to voiced stops on a number of acoustic variables, providing a baseline against ‘nasal+voiceless stop’ sequences can be compared. Second, it will examine the effect boundary strength exercises on voiceless and voiced stops,

again providing an understanding of the type of boundary strength effects in the absence of the sandhi site, and giving baselines for forming expectations as to how boundary strength effects will be realised on ‘nasal+voiceless stop’ sequences if there is a categorical or a variable boundary strength effect.

3. **A production experiment on ‘nasal+voiceless stop’ sequences** and the effect boundary strength exercises on the measured variables separating voiceless from voiced stops.

This is the main experiment in which speakers produce ‘nasal+voiceless stop’ sequences. The effect of boundary strength is measured on a number of variables, testing whether it is qualitative or variable, concluding on whether there is support for the existence of labelled domains, as Nespor and Vogel proposed.

In the course of the analysis first we present the production experiment on voiceless and voiced stops, since it is the baseline against which all following results will be compared. Then, results from the pilot perception experiment are discussed, giving a first insight to whether the phenomenon is a sandhi or coarticulatory process. Finally, on the basis of baseline measurements we form expectations for the acoustic realisation of ‘nasal+stop sequences’, which are subsequently presented, concluding as to whether there is a categorical effect on the process or not.

4.2 Methodology

First the methodology for the production experiment is presented because the perception experiment was based on the production one. Listeners were asked to classify whether voicing occurred on the items produced during the production experiment.

4.2.1 Production experiment

4.2.1.1 Materials

As mentioned in Chapter 2, the template used for constructing items was ‘ADJECTIVE + NOUN’. For the purposes of this experiment the Adjective is expected to end with a nasal, and the Noun to begin with a voiceless stop. In Greek, of the

1	[a'ðapanon ta'belon] 'free signs'	2	[pa'ranomon pi'lonon] 'illegal polls'
3	[pti'somenon pa'ceton] 'foldable packs'	4	[a'cinðinon pa'rasiton] 'harmless parasites'
5	[a'prozmenon ci'tiðon] 'unexpected starting points'	6	[xar'mosinon ka'banon] 'happy-sounding bells'
7	[a'ðinamon ti'fonon] 'weak typhoon'	8	[a'cinðinon ka'zino] 'harmless casinos'

Table 4.2: The 8 items constructed for investigating SV.

two nasal consonants only /n/ is allowed word finally², thus all Adjectives in this design end in /n/. The genitive of plural of most adjectives in Greek ends in /-on/. For this experiment adjectives whose last syllable in the genitive of plural is either /-non/ or /-mon/ were specifically chosen with the purpose of keeping the pre-boundary vowel constant³. Example 4.3 shows an item constructed for this set, where stop voicing could occur between the Adjective and the Noun.

(4.3) [pa'ranomon **pi'**lonon] 'illegal polls' - GEN PLURAL, MASC

Items containing all three voiceless stops were created, as can be seen from the list of items constructed (Table 4.2). This set of items meant to test the application of stop voicing will be referred to as 'nasal+stop' sequences representing the ambiguous items where stop voicing might occur or not.

Added to this set of materials was also one final condition, anticipated to represent the lowest prosodic condition. It was made up using the template presented in Chapter 2: 'definite article + verb'. However, for this experiment solely, the pre-boundary word is not only a definite article but a variety of lexical items, as can be seen in Table 4.3. The reason for incorporating temporal, subjunctive and negation particles in this analysis is testing Nespor & Vogel's (1986) statement that the

²Sometimes /m/ can also be found in loan words, such as ['tram].

³For amplitude normalisation purposes, amplitude measurements are taken from each vowel and are subtracted from the amplitude measurements taken from each stop. Since vowels have different intrinsic amplitudes, it was important to have the same vowel in all conditions and items for all comparisons. This is further explained in the Measurements section.

MATERIALS WITH PARTICLES	
PARTICLE	VERB
['otan] 'when'	[pace'taris] 'pack'
[an] 'if'	[taksi'ðevis] 'travel'
[ðen] 'not'	[kata'gelis] 'denounce'
[min] 'not'	
[ton] 'the'	

Table 4.3: The 15 items constructed for measuring stop voicing in the lowest condition (BC0).

1	[a'siðota pa'zarja] 'hard bargains'	['omorfa ba'gazja] 'beautiful luggage'
2	[epi'cinðina ta'ksiðja] 'dangerous travels'	['evixa da'uða] 'nice-sounding drums'
3	[a'ðjakopa ka'lezmata] 'constant callings'	['eksipna ga'rsopa] 'clever waiters'

Table 4.4: The 3 items constructed using voiceless and voiced stops not preceded by a nasal for baseline measurements. These items were embedded in the same conditions as their voiceless counterparts. The items for measuring voiceless stops were taken from the design for pre-boundary lengthening.

temporal particle ['otan] blocks stop voicing forming a clitic group of its own, while the negation particle [ðen] does not.

For the experiment testing the acoustic differences between voiceless and voiced stops, baseline conditions were also created and recorded, against which the stop voicing conditions are compared, and which provide insights into how boundary strength is manifested in phonologically voiceless and voiced stops. Six items were constructed, three containing voiceless stops and three containing voiced stops, and they were embedded in sentences representing boundary conditions BC1-BC5. Table 4.4 presents the items constructed to act as baselines. Naturally, no preceding nasal is used in these conditions, and the template for the pre- and post-boundary words remains 'Adjective+Noun'. Table 4.5 presents a breakdown of the number of items used for this experiment.

		Speakers	Items	BC	Repetitions	Particles	TOTAL
SV	Adj+Noun	5	8	5	2		400
	Part+Verb	5	3	1	2	5	150
							550
Baselines	Voiceless	5	3	5	2		150
	Voiced	5	3	5	2		150
							300

Table 4.5: The number of items constructed for measuring stop voicing (SV), and the number of items for baselines, broken down by number of speakers, items, boundary conditions, repetitions, and the number of particles used.

4.2.1.2 Segmentation

Where possible, a complete segmentation was carried out for every segment of the pre- and post-boundary syllables. Measurements were taken from the closure and the release of the post-boundary stops. The beginning of the closure was marked when there was a clear change in the waveform moving from the voiced nasal to the voiceless closure, combined with the end of the F2 formant, if possible. The beginning of the release was marked where there was again a clear amplitude change in the waveform, and the end of the release was signalled by the first periodic cycle and the clear existence of voicing-excited formants. Only the release is marked, since voiced stops in Greek are fully prevoiced, and many times it is also not possible to distinguish carryover voicing from negative VOT, while the release offers more clear and easily identifiable landmarks for segmentation. Figure 4.1 shows a typical example of segmentation, with the important landmarks segmented; the pre-boundary vowel (pV) and nasal (N), and the post-boundary stop (closure ‘C’ and Release) and vowel (fV). Some times it was also easy to segment the onset /n/ or /m/ of the pre-boundary syllable (pre-boundary syllables comprised of either /mon/ or /non/).

4.2.1.3 Measurements

1. Duration of Closure and Release

The duration of the closure and of the release are known to be significant fac-

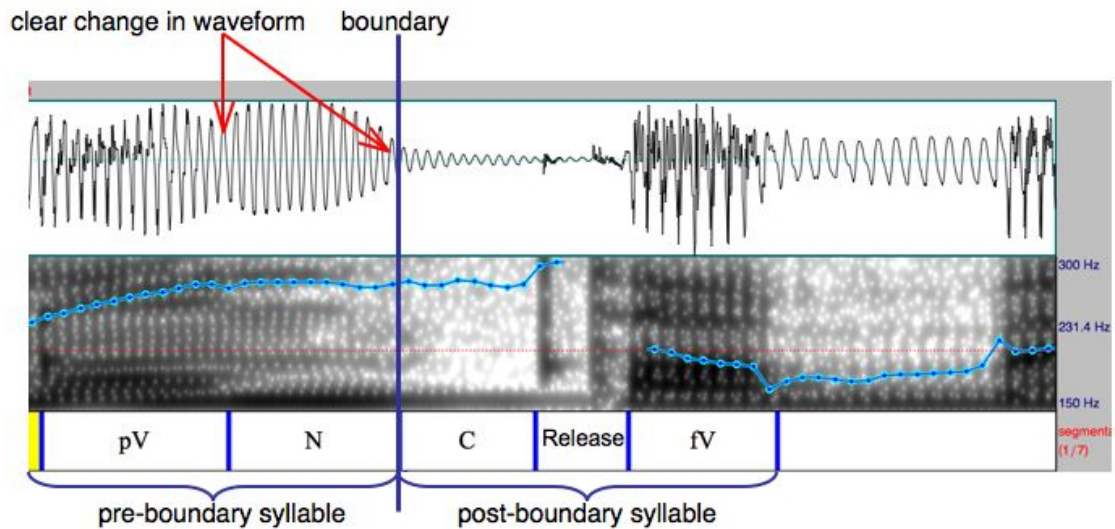


Figure 4.1: Example of segmentation of items constructed for the investigation of post-nasal stop-voicing. Sample taken from a female speaker saying the item [a'ðapanon ta'belon] (romanisation: *adapanwn tampelwn*). Figure shows the names and positions of the syllables, and of their segmentation. It also points to the position of the boundary, as well as to positions that were important for segmentation decisions. This example is taken from BC2.

tors distinguishing voiceless and voiced stops in other languages (Port 1979), and are expected to act so in Modern Greek too.

2. Relative Amplitude of Closure and Release

The amplitude of the closure and release have been found in other languages to separate between the two categories of sounds (Lisker & Abramson 1964, Lisker & Abramson 1967, Hayes & Stivers 2000, Cole et al. 2007, Kong et al. 2007). For the current experiment, the amplitude at the mid-point of the closure and of the release was measured for each rendition. Many different measurements have been employed regarding amplitude by different studies in the past, depending on the types of sounds investigated. For example, the studies mentioned by Lisker included nasal sounds, which were compared to the voiced, voiceless and prenasalised sounds using a combination of measurements within the closure of the consonant (Kong 2007 also addressed the issue of nasality differences across voicing contrasts using amplitude measurements).

For the present experiment, only the mid-point of the closure and the release were selected as the most representative points, and also as a way of maintaining a small number of dependent variables. Figure 4.2 shows the window over which PRAAT was set to measure amplitude. Amplitude was automatically measured by Praat, but the window over which intensity is measured relies on the “minimum pitch” value provided by the user. The minimum pitch for this experiment was set at 100Hz, meaning that the resulting window frame over which intensity is calculated was 18ms (9ms on each side of the point of measurement, which was placed at the midpoint of the closure and release respectively).

Amplitude measurements can be affected by a number of things (like, for example, how close the speaker is to the microphone). Therefore, for normalisation purposes, the amplitude of the closure and the release was measured *relative* to that of the pre-boundary vowel. The pre-boundary vowel was always an /o/ therefore to avoid confounds in the measurements; vowels can have different intrinsic amplitudes, which would interfere with the relative measurements of amplitude. Thus, the mid-point amplitude measurement of the preboundary vowel /o/ was subtracted from the midpoint amplitude measurement of the closure and the release:

$$\text{Relative Amplitude} = \text{dB of Closure/Release (midpoint)} - \text{dB of Vowel (midpoint)} \quad (4.4)$$

Since vowels generally have higher amplitude than the closure of stops negative values are expected from the application of the above equation. Lower absolute values, therefore, indicate higher amplitude (e.g. a value of -8 means that the closure has more amplitude than a value of e.g. -11).

Although there is a number of different measurements that could have been taken for voiceless - voiced stop distinctions, several were discarded for not being applicable to the data. For example, F0 has been known to distinguish the voicing contrasts; Stevens (1998, pp. 476) mentions that the F0 at the onset of the vowel immediately following a voiced stop is lowered by 5 to 7 percent in comparison to the voiceless stop. However, given that the F0 patterns varied across conditions,

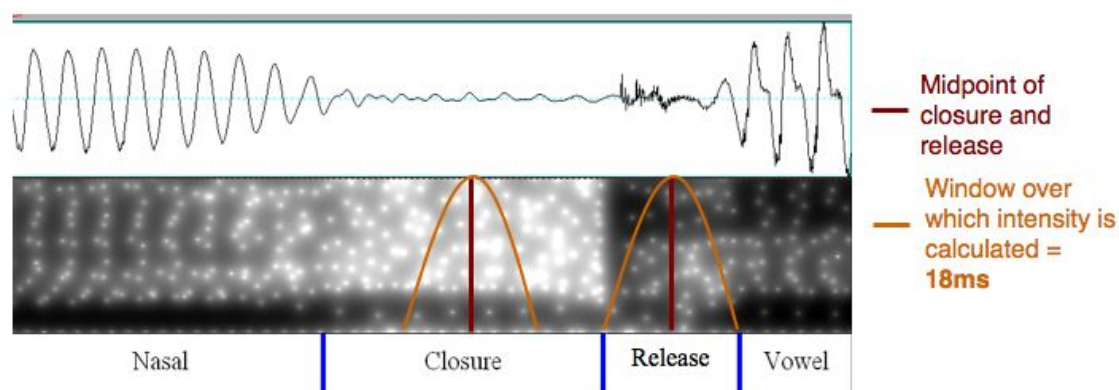


Figure 4.2: Example of how the relative amplitude of the closure and the burst were measured. This figure shows the window over which the amplitude was measured using PRAAT.

and they constitute a matter of investigation in this design, it was not possible to use this as a means of voicing identification.

Hence, in all analyses in this chapter, the dependent variables will be the duration and relative amplitude of the closure and release.

Distribution of deletion of preboundary nasal

One final extra measurement was performed; although in most instances the pre-boundary nasal could be seen and segmented, there were cases where the nasal seemed to have been deleted. This was in accord with Arvaniti's (2007) suggestion that the nasal contributing to the voicing can be completely deleted. The instances where there was no trace of the nasal were marked and used as an added means of testing whether the effect of boundary strength is categorical or not by testing how many times deletion is found to occur across boundary conditions. Additionally, these instances were also marked when perceptually there was no nasal preceding the stop, therefore this is an impressionistic analysis with some phonetic support, but it has only been performed as a preliminary analysis of whether nasal deletion co-occurs with voicing, and on which further acoustic measurements of nasality can be performed in future experiments.

4.2.2 Preliminary perception experiment

4.2.2.1 Design

Three raters were asked to classify all produced items; both ‘nasal+stop’ sequences and the baselines. The latter were added to make sure that they are recognised with high accuracy, since they are the unambiguous phonological cases. All three listeners were in their late twenties, and were native speakers of Greek. Two were expert phoneticians, and one was a ‘naive’ listener.

Given that the materials presented to the listeners were extracted from the production database, it is obvious that the stop consonant was always situated at the beginning of a word. As has been presented already, the word ending in a nasal was always an Adjective ending in a nasal, and the word starting with the stop was a Noun; e.g. [a'cinðinon ka'zino]. For the lowest condition, the word ending in a nasal was either a definite article (e.g. [ton]), a temporal or conditional particle (e.g. ['otan], [an]), or a negative particle (e.g. [ðen], [min]), and the word starting with a stop was a Verb. The sound played back to the rater only consisted of the syllables preceding and following the word boundary. An example of segmentation of an instance of a voiced bilabial stop can be found in Figure 4.4. It was decided against playing the whole words, since lexical bias is to be avoided when making such judgements. Three sets of wav files were created, one for each place of articulation of the stop consonants of Modern Greek. Each rater heard the same number of sounds within each set, but they were randomised differently. There were 272 wav files for bilabials, 268 for velars, and 206 for alveolar stops.

The experiment was run in PRAAT Boersma (2001), using an adjusted version of the MFC Identification script provided by the software. A screenshot of the graphical user interface used for the experiment is presented in Figure 4.3. The rater could choose between four options; a voiced or voiceless segment with the subcategorisation of it being prenasalised or not. E.g. for bilabials the available options were /p/, /np/, /b/, /nb/. Given that the main experimental interest resided in the distinction between voiced and voiceless sounds, those four categories were later collapsed to two for the purposes of the analyses; the instances classified as a voiceless stop (preceded or not by a nasal) were considered as a voiceless stop, and the

ones that were classified as voiced were regarded as such irrelevantly of whether a nasal preceded or not. A Replay and a No Answer (NA) button was provided, and the rater was allowed to listen to the sound up to three times if needed to arrive to a conclusion. Finally, an ‘Oops’ button was provided, in case the rater made a mistake. Each rater did the experiment three times - once for each set of stops.

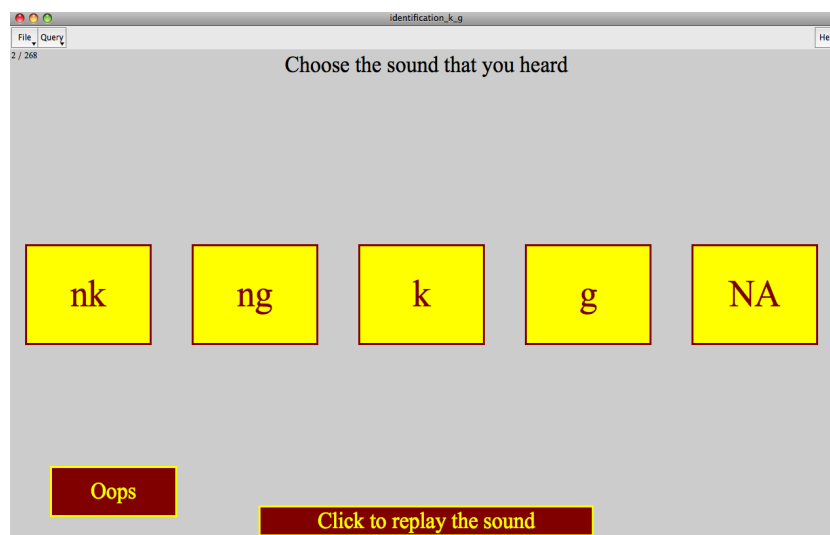


Figure 4.3: Example of the screen the raters were seeing when running the experiment. Example taken from data for /nk/ classification. Instances classified as either /k/ or /nk/ were later collapsed to a category called ‘voiceless’, while instances classified as either /ng/ or /g/ were collapsed to a category called ‘voiced’. Same categorisation took place for the other two types of stops (/t/ and /p/). The No Answer, Replay, and Oops buttons are seen.

4.2.2.2 Measuring inter-transcriber agreement

Agreement with *Kappa* Statistic

Before analysing the perceptual data from the three raters, it is important to quantify whether the raters agree in their judgements, and to what extent. In order to statistically evaluate the agreement between all three raters, the *kappa* statistic was used (Siegel & Castellan 1988), computed as shown in equation (4.5). This statistical measure takes into consideration the probability of chance agreement between raters, and outputs have a value between -1 (no agreement between raters) and 1 (full agreement) (see e.g. Carletta 1996, for an analysis of why kappa is important). Therefore, expected results should vary between 0 and 1 with 0 being the point where raters are expected to agree by chance.

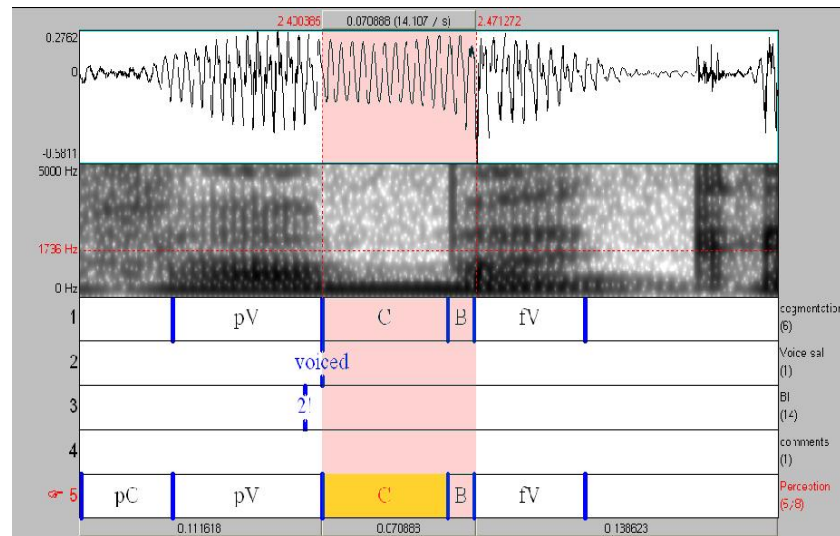


Figure 4.4: Example of the part extracted from the waveform. The shaded part is the target consonant. In Tier 5 the segmentation from pC to fV shows the part of the sound that was extracted and played to the rater for identification.

$$k = \frac{P(A) - P(E)}{1 - P(E)} \quad (4.5)$$

Siegel & Castellan’s (1988) definition of the kappa statistic is:

‘The kappa coefficient of agreement is the ratio of the proportion of times that the raters agree (corrected for chance agreement) to the maximum proportion of times that the raters could agree (corrected for agreement)[...] $P(A)$ is the proportion of times that the κ raters agree and $P(E)$ is the proportion of times that we would expect the κ raters to agree by chance.’ (pp.285)⁴

⁴The step for calculating $P(E)$ is described in Siegel and Castellan (1988, pp. 290) as “For each category j [each category is the possible outcomes available to the raters] find the number of times that any object is assigned to that category. This number is C_j . Next find p_j , the proportion of rating assigned to the j th category. Then, by using equation

$$P(E) = \sum_{j=1}^m p_j^2 \quad (4.6)$$

While it is a widely used statistic, and is generally considered to be a much better indication of agreement than raw percentages, it is not generally agreed upon which levels can be considered as *acceptable*. Landis & Koch (1977, pp. 165) propose as “benchmarks” that results between 0.41-0.60 represent ‘Moderate’ agreement, 0.61-0.80 ‘Substantial’ agreement, and 0.81-1.00 ‘Almost Perfect’. Within the field of Computational Linguistics, it is generally accepted that a kappa statistic above 0.65 can be accepted for further analyses of the data (personal communication with Dr. Jean Carletta). Therefore, it is clear that there isn’t much consensus regarding the exact use of the statistic, but it is generally agreed that a result above 0.70 represents an acceptable level of inter-transcriber agreement.

4.2.2.3 Analyses

The preliminary perception experiment will be analysed by measuring how many times raters classified ‘nasal+voiceless stop’ sequences as being voiced or voiceless across boundary conditions. This classification will also be used later on to test whether the acoustic manifestation of the two categories is categorically distinct, or whether it varies on a continuum of changes, in order to conclude whether the phenomenon appears to be a sandhi process or the result of coarticulation.

4.3 Results

4.3.1 Production experiment on voiceless and voiced stops

A study testing the phonetic differences between phonologically voiceless and voiced stops is performed to identify (a) the correlates that distinguish voiceless from voiced stops, and (b) the effect that boundary strength exercises separately on each category. The duration and relative amplitude of the stops’ closure and release are measured. First the acoustic differences between the two phonological categories are identified, and using a regression and a discriminant analysis the phonetic correlates that are most important for signalling the voicing distinction are identified. Then, the effect of boundary strength on each set (voiceless vs voiced stops) is measured.

find $P(E)$, the expected proportion of agreement among raters had they been rating the objects at random.”

		Voiceless	Voiced	Significance
Closure Duration	ms	51	64	*<.001
Release Duration		24	14	*<.001
Closure Amplitude	dB	-17	-3	*<.001
Release Amplitude		-10	-2	*<.001

Table 4.6: Mean duration (ms) and amplitude (dB) of the closure and the release of phonologically voiceless and voiced stops, as produced by the 5 speakers.

These results provide hypotheses for how boundary strength will affect the phonetic manifestation of ‘nasal+stop’ sequences in the following section.

4.3.1.1 Acoustic correlates of voicing distinctions

Three items with a voiceless and three with a voiced stop in post-boundary condition were placed in conditions BC1-BC5 for the purposes of this experiment, and produced by 5 speakers (two repetitions of each - Table 4.4). Four sets of Univariate ANOVAs were performed, one for each dependent variable; duration and relative amplitude of closure and release, with voicing variation as the independent variable with two levels; voiceless and voiced.

Initially all places of articulation (velar, alveolar and bilabial) and all boundary conditions are pooled together. All dependent variables are significantly different between the voiceless and the voiced renditions. Table 4.6 shows the mean values for each variable. The closure of the voiceless renditions was significantly shorter than the voiced ones, with voiceless closure being 64ms and the voiced 51ms, therefore 20% shorter ($F(1,271)=45.521$, $p<.001$). The duration of the release was significantly longer for the voiceless stops, 24ms vs 14ms, i.e. 42% longer ($F(1,285)=150.290$, $p<.001$). The amplitude of the closure and of the release were significantly lower for the voiceless than for the voiced renditions. The amplitude of the voiceless closure was -17dB when the voiced was -3dB, a difference of 82%, and the amplitude of the voiceless release was -10dB vs -2dB for the voiced, a difference of 80% ($F(1,247)=563.993$, $p<.001$, and $F(1,258)=295.8$, $p<.001$ respectively)⁵.

⁵The same analyses were performed for each place of articulation separately, and the voicing distinction was significant for all measurements in all places of articulation (except for the duration

All measured variables appear, therefore, to distinguish statistically between the voicing contrast. A Binary Logistic Regression analysis was performed with the duration and amplitude of the closure and release as the four predictor values predicting whether each item is likely to be a voiceless or voiced stop. It was found that the acoustic characteristics of the *closure* were the most important predictors contributing to the classification of the items, while those of the burst did not significantly contribute to the model (see Table C.3 in Appendix C for results). A discriminant analysis was then performed using only the duration and amplitude of the closure to see how well the two variables predicted the voicing contrast, and the result showed that 98.4% of all baseline instances can be correctly classified based on the acoustic characteristics of the closure. This suggests that the closure is the most important cue in signalling the voicing contrast in Modern Greek. That is not to say that the release might not be significant in the voicing manifestation. First, the duration and amplitude of the burst were significantly different across voiceless and voiced renditions; second, the regression model specified that any of the four variables could significantly change the performance of the model but when added to the model, only the duration and amplitude of closure were significant; and, third, a discriminant analysis testing the predictive power of the acoustic characteristics of the release classified 88.1% of the instances correctly suggesting that the release too must carry information which might not be as marked as that of the closure.

4.3.1.2 Boundary strength effect on voiceless and voiced stops

The previous section established that phonological voicing distinctions are signalled via all four acoustic correlates measured, with those of the closure being the most important. The next question tests the effect boundary strength exercises on post-boundary voiceless and voiced stops. It was hinted in Chapter 3 that in post-boundary positions there are no clear duration effects either on the closure or on the burst of the stop other than a non-significant tendency for some lengthening. The present section investigates the effect of boundary strength on voiceless and voiced stops on all four variables (duration and amplitude of closure and release).

of the closure of the bilabial sounds - see Table C.1 in Appendix C for results, mean values for each variable, and statistics for each place of articulation).

		Voiceless						
		BC1	BC2	BC3	BC4	BC5	Sign	Bonferroni
Duration (ms)	Closure	45	49	52	54	60	.016	1,2<5
	Release	22	24	24	24	25	-	-
Amplitude (dB)	Closure	-13	-13	-16	-17	-21	<.001	1,2,3,4<5
	Release	-8	-11	-11	-11	-7	-	-
		Voiced						
Duration (ms)	Closure	56	60	66	70	69	.002	1<4,5
	Release	13	14	14	13	14	-	-
Amplitude (dB)	Closure	-3	-3	-4	-4	-1	<.001	1,2,3,4<5
	Release	-3	-2	-2	-2	1	.013	1,2,3,4<5

Table 4.7: Mean duration (ms) and amplitude (dB) of the closure and the release of phonologically voiceless and voiced stops, as produced by the 5 speakers for the pilot measurements.

Four REPEATED MEASURES ANOVAs were performed on each variable with boundary strength as the independent variable (5 levels).

Boundary strength significantly affected the duration of all variables except for the release of voiced stops (VOICELESS STOPS: Duration Closure $F(4, 36) = 4.151$, $p=.007$, Duration Release $F(4, 84) = 3.5$, $p=.011$, VOICED STOPS: Duration Closure $F(4, 88) = 8.5$, $p<.001$, Duration Release $F<1$). Also, boundary strength affected the amplitude of all variables except for the release of voiceless stops (VOICELESS STOPS: Amplitude Closure $F(4, 44) = 12.96$, $p<.001$, Amplitude Release $F<1$, VOICED STOPS: Amplitude Closure $F(4, 84) = 6.7$, $p<.001$, Amplitude Release $F(4, 36) = 3.7$, $p=.013$). Figure 4.5 shows the effect of boundary strength on the four dependent variables broken down by voicing condition. The left panel shows the amplitude measurements, and the right panel shows the duration measurements, and the left figure within each panel shows the voiceless, while the right figure the voiced items. The results are also presented in Table 4.7, where Bonferroni post-hoc analyses show how exactly the effect of boundary strength takes place across boundary conditions.

All results are in the anticipated direction; any boundary strength effects apparent on duration show lengthening, and any effects on amplitude show that there is lower amplitude higher in the structure. Especially regarding amplitude measurements,

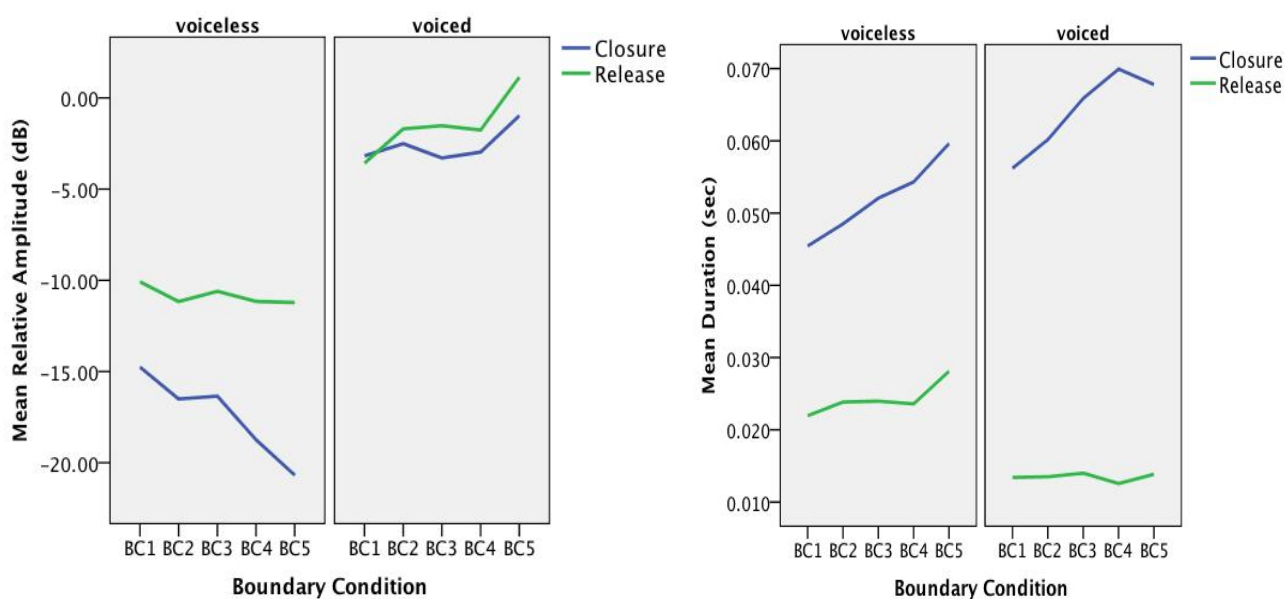


Figure 4.5: Effect of boundary condition on the duration and amplitude of the closure and the release of the phonologically voiceless and voiced renditions.

the effects are very small even where significant differences do arise. In general, the acoustic characteristics of the release do not seem to be affected significantly by boundary strength; even for the amplitude of the release of voiced stops, for which a significant effect arose, it was due to a significant difference of BC5, the only condition with most missing items due to pauses. Therefore, the release of voiceless and voiced stops does not seem to be influenced significantly by boundary strength.

The closure of both voiceless and voiced stops shows a tendency for lengthening and lower amplitude in higher domains, and any boundary strength effects are more pronounced in the closure. Almost all effects found are derived from higher domains being different to lower domains, with variable effects in between. We take these results to suggest that the closure of post-boundary stops is most affected by boundary strength, that the effect is variable across conditions, and that is not as pronounced as it was seen in pre-boundary positions in Chapter 3, supporting the findings of that chapter.

In conclusion, the following general observations can be deduced from this experiment:

1. All acoustic correlates measured were statistically significant for marking voicing distinctions: the duration of voiceless closure is significantly shorter than for voiced stops, voiceless release is significantly longer than voiced, and the amplitude of both closure and release is lower in voiceless than voiced sounds. A regression model showed that mainly the acoustic characteristics of the closure are important in signalling the voicing differences, but the duration and amplitude of the release are also important in modelling voicing distinctions in the absence of closure information.
2. The boundary strength effect is most prominent on the phonetic characteristics of the closure again with long durations and lower amplitude in high domains, and in that the closure exhibits the gradual effect expected by boundary strength, as it was shown in the previous chapter. The effect is not very pronounced, and seems to be a *trend* for lengthening and lowering of the amplitude at stronger boundaries rather than a clear significant effect. It is taken to support the finding from Chapter 3, that post-boundary effects do not clearly exist in Modern Greek. In the few instances where an effect was apparent, like in the duration of the closure, the resulting values do not create ambiguities across voicing contrasts since values do not overlap (e.g. the lowest duration of voiceless release was 22ms, and the highest of voiced release was 14ms - with the exception of the duration of the closure of the voiced stops overlapping with that of the voiceless across BC0 and BC4,BC5).

Based on the findings just discussed it is now possible to continue with investigating ‘nasal+voiceless stop’ sequences.

4.3.2 Preliminary perception experiment

Having identified the nature of the phonological distinction of voicing between voiceless and voiced stops in Modern Greek, we move on to the investigation of the ambiguous ‘nasal+stop’ sequences, that is the sites which can give rise to post-nasal stop voicing. A preliminary perception experiment is first performed to test the perceptual validity of Nespor & Vogel’s (1986) impressionistic analyses of post-nasal

stop voicing in Modern Greek, and subsequently to identify whether perceptually the phenomenon seems to be influenced by boundary strength in a qualitative manner.

Following the perceptual investigation of post-nasal stop voicing, phonetic analysis on the data investigates the question of whether the nature of the process is phonetic or phonological. By investigating the phonetic characteristics of instances that were classified as voiceless vs voiced, it is possible to conclude whether the phenomenon is acoustically gradient or categorical. Following that is the main investigation of the effect of boundary strength on the phonetic manifestation of the process, and therefore of the distribution of stop voicing occurrences across prosodic domains. Based on the findings from the first experiment on voiceless vs voiced stop acoustic differences, hypotheses are formed as to what the phonetic manifestation of each acoustic correlate will look like if the phenomenon occurs categorically or gradiently across boundary conditions, and therefore conclusions are drawn that can then be compared to Nespor & Vogel's (1986) analysis, and which can also answer the question of whether qualitative distinctions between prosodic domains can be identified on the basis of such sandhi-type phenomena.

4.3.2.1 Inter-transcriber agreement

Before discussing the results, the kappa statistic was calculated to confirm that raters agreed in their perception of when stop voicing occurs and when not. The statistic was calculated three times, once for each set of stops (i.e. alveolar, bilabial and velar). The raters' agreement was very high for all three sets: $\kappa = 0.888$ for alveolars, $\kappa = 0.845$ for bilabials, and finally $\kappa = 0.901$ for velars (close to values of 1, signalling full agreement).

This suggests that voicing distinctions are easy to detect reliably. Table 4.8 shows how many times (percent) each rater identified each stop as voiced or voiceless⁶. The high percentages in the voiced segments recognised as voiced show that the task was easy for the raters. In the columns for voiceless segments recognised as voiced, one can see that there are quite higher percentages of *voiceless stops classified as voiced*,

⁶Sometimes the percentages do not add up to 100%. This is due to the option the raters had of saying 'No Answer', if they found it too difficult to make a decision.

indicating that indeed there must exist instances of stop voicing in the database.

Rater	EK		AL		GP	
	Voiceless	Voiced	Voiceless	Voiced	Voiceless	Voiced
np	71.6	24.8	72	25	65.7	34
(n)b	1	99	5.1	94.9	0	100
nk	73.5	22.2	84.7	15.3	65.9	33.7
(n)g	1	97	2.3	97.7	1.1	98.9
nt	61.02	38.9	56.9	43.1	52.2	47.7
(n)d	1.5	98.5	3	97	0	100

Table 4.8: Results in percentages for each rater. Rows represent the percentage that each segment was classified as a voiced or voiceless stop. Some small percentages might not be adding up to 100% given that the option of ‘No Answer’ was offered to the raters. Numbers in bold indicate instances of voiceless stops classified as voiced.

These results highlight two issues: first, there are percepts of voicing in the data, as seen from the fact that underlyingly voiceless items have been classified as voiced. Second, high inter-transcriber agreement suggests that the difference between instances of ‘nasal+stop’ sequences that have and haven’t undergone voicing must be easy to detect, and therefore would be expected to be phonetically marked. It is to that data that we now turn, to see how they are distributed across boundary conditions.

4.3.2.2 Prosodic Boundary Effect

In this section, only the ‘nasal+stop’ sequences are analysed. Table 4.9 and Figure 4.3 show the percentage of times ‘nasal+stop’ sequences were classified as voiced in each prosodic condition, pooling all raters together. It is clear that most voicing instances occur within BC0; 72.5% of all /n#t/ instances in BC0 are classified as voiced, 51.1% of /n#p/ and 51.6% of /n#k/, while there is a radical drop in the number of voiced instances in all other conditions⁷. Figure 4.6 shows the distribution of ‘nasal+stop’ sequences classified as voiceless and voiced across conditions pooling across place of articulation.

⁷Note that the n#t sequences have more instances of voicing, which could be due to /n/ and /t/ being homorganic. This could suggest that articulation is an important factor in the output of stop-voicing within the condition where voicing takes place.

	% Classification as voiced		
	nt	np	nk
BC0	77%	56%	54%
BC1	25%	4%	0%
BC2	16%	0%	0%
BC3	19%	0%	0%
BC4	0%	0%	0%
BC5	0%	0%	0%

Table 4.9: Results in percentages for classification of voiceless stops as voiced by all raters for each boundary condition. The percentage for the first boundary condition is higher, suggesting that stop-voicing mainly occurs on this level.

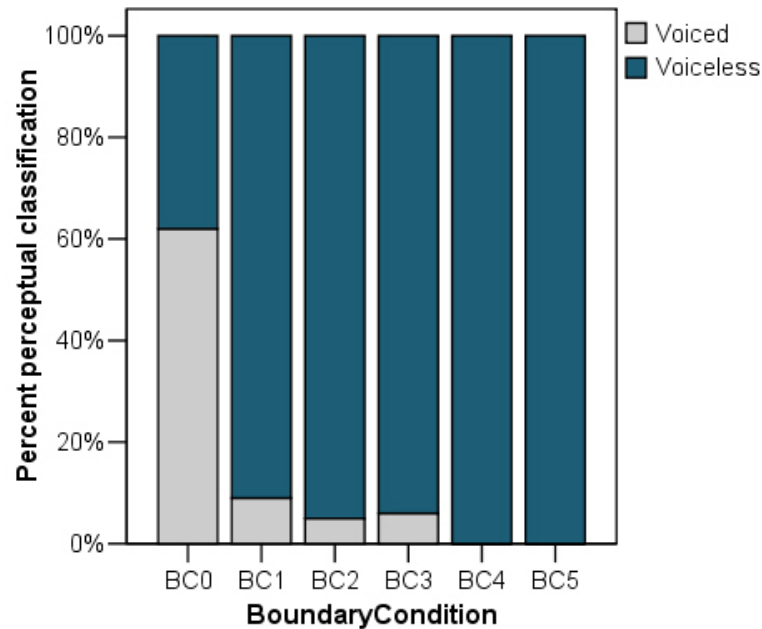


Figure 4.6: Percentage of perceptual classification of voiceless instances (all places of articulation and speakers grouped) as voiced or voiceless in each prosodic condition.

The fact that the majority of instances of stop voicing is perceived to occur within BC0 supports Nespor & Vogel’s (1986) analysis that there is a qualitative difference across lower and higher domains in the way post-nasal stop voicing takes place. Nespor & Vogel (1986) also discussed differences across lexical items, like *όταν* ‘when’ vs *δεν* ‘not’. These are discussed in the following section.

4.3.2.3 Lexical differences in the lowest domain

When designing BC0 for the present experiment, the items created entailed a variety of lexical items ranging from negation and temporal particles to definite articles. The reason for incorporating that in the design was to further probe Nespor & Vogel’s (1986) impressionistic analysis, which proposed that there are differences across lexical items in the way they relate to prosodic constituency. For example, they mentioned that the words [’otan]+Verb form two prosodic domains, while [ðen]+Verb are one (i.e. temporal vs negation particles), and this provided support for proposing the existence of the clitic group (although their proposal does not refer to them as different lexical items, and it does not provide an explanation why the temporal particle will have a different prosodic status to the negative particle).

Item	Agreement		
	Voiceless	Voiced	No Agreement
’otan	10	3	7
ðen	-	19	5
an	12	5	5
min	8	11	9
ton/tin	5	17	5

Table 4.10: Differences between lexical items as to how many were classified as voiced, voiceless, or didn’t reach agreement.

Breaking down the instances classified as voiceless and voiced by lexical item the result (shown in Table 4.10) supports Nespor and Vogel’s perceptual classifications; the negation particle [ðen] voiced 100% of the times, while the temporal particle [’otan] only voiced 23% of the times (3 out of 13 times that reached agreement across raters). The pattern seen in the temporal particle [’otan] is replicated with the subjunctive particle [an], while negative particle [min] and definite articles [tin/ton] seem to group with [ðen], but not completely, since they also have voiceless and no agreement instances.

Therefore, while Nespor and Vogel’s analysis of difference between [’otan] and [ðen] was replicated in the present data, their statement that there is a qualitative distinction between prosodic contexts where voicing is allowed and where it isn’t is not supported by the present results. There is a number of issues that go against

their expectations, that need to be pointed out, and that suggest that there might be both coarticulation in the application of stop voicing, and variable differences across lexical items.

First, there is a number of instances that did not reach agreement across raters. These were further analysed statistically, and are discussed in section 4.3.3.1. Second, the instances that didn't reach agreement were not due to only one speaker. 'No agreement' classifications had been given for all 5 speakers, showing that it is not only one speaker that might be producing instances in a different manner. One of five speakers voiced almost all instances in BC0, suggesting that there are differences across speakers, like the one proposed in Ellis & Hardcastle (2002). All other speakers had voiced and voiceless items. Third, the particles [min], [an] and definite articles [ton/tin] which were not investigated in Nespors and Vogel show that the distribution of how many times voicing takes place within BC0 cannot be fitted into neatly distinct categories of where the phenomenon is allowed and where it is blocked, as would be their expectation.

All the above show that, while the phenomenon of stop voicing perceptually is contained within the lowest domain suggesting a qualitative boundary strength effect, there is variation on its application within this domain. This variation is seen in two ways; first, some aspects of the phonetic manifestation of stops seem to create ambiguity in that transcribers do not reach complete agreement. This suggests that there might exist some type of phonetically gradient output which might have not been captured in the present design, since we have not investigated the behaviour of the pre-boundary nasal, and second, there are differences across lexical items on how many times they voice within BC0, with some items voicing more and some fewer times.

4.3.2.4 Conclusions on perception test

The perceptual analysis of post-nasal stop voicing in this section has shown that the traditional impressionistic analysis by Nespors & Vogel (1986) is replicated using a perception experiment, supporting their claim that boundary strength affects the application of post-nasal stop voicing in a qualitative manner, whereby stop voicing is optional in the lowest domain, and categorically blocked in all higher domains, and

whereby there is a distinct difference between [ˈotan] and [ðen]. However, it has also provided evidence that Nespov and Vogel did not capture the variability that exists within the lowest domains; they tried to capture differences across lexical items by proposing the existence of a further prosodic domain (i.e. the clitic group), however perceptual evidence shows that the distribution of how many times voicing takes place within and across lexical items contains much variability. Importantly, the present experiment has provided a separation of the ‘nasal+stop’ data in instances perceived as voiced vs voiceless, a distinction useful for identifying the acoustic differences between the two categories, therefore allowing insights as to whether a categorically perceived process is phonetically gradient or not.

4.3.3 Production experiment of ‘nasal+voiceless stop’ sequences

The goal of this section is twofold; first, to establish whether the application of stop voicing is phonetically gradient or categorical. That is, whether the acoustic values distinguishing ‘nasal+stop’ sequences classified as voiceless vs voiced range in a continuum of possible phonetic outputs, or whether they are grouped categorically in two distinct distributions potentially forming two distinct phonological domains. Second, it is the goal of this section to identify the effect of boundary strength on each phonetic variable measured, and to conclude whether the effect exercised by boundary strength on the application of stop voicing supports an analysis of qualitatively distinct domains, or whether it is variable.

Both questions have been tentatively answered by the perception experiment, which suggested that the process seems to be phonological, but there might be coarticulation or some type of variation within prosodic domains where it occurs, and its application seems to be qualitatively influenced by boundary strength. The phonetic analysis in the present section will provide insights into how the application of the phenomenon interacts with boundary strength effects, and, importantly, whether the phonetic data support the perceptual analysis.

4.3.3.1 Sandhi or coarticulation?

As explained in the Introduction, it can be postulated that a sandhi-type phenomenon will take place categorically, while a coarticulatory process gradiently, and

	Duration		Amplitude	
	Closure	Release	Closure	Release
voiceless	43	24	-14	-18
voiced	38	15	-8	-7

Table 4.11: Mean duration (ms) and amplitude (dB) values for all ‘nasal+voiceless stop’ sequences classified as voiceless and voiced by raters.

its application cannot be predicted by a set of traditional phonological rules. In the present section the phonetic manifestation of ‘nasal+stop’ sequences is investigated to test whether the phonetic characteristics of instances classified as voiceless are categorically distinct from those classified as voiced.

The experiment which tested the acoustic differences between phonologically voiceless and voiced stops reported that voiceless stops had significantly shorter closure durations than voiced stops, longer release durations, and smaller amplitude for closure and release. This section tests whether this pattern is replicated on ‘nasal+voiceless stop’ sequences that were *classified* by the raters as voiceless vs voiced, or whether the differences across the two classifications are variably ranging on the phonetic continuum.

The pattern of voiceless vs voiced differences was replicated on the instances that were classified by the raters as voiced vs voiceless: four REPEATED MEASURES ANOVAs were run, one for each of the variables measured, with voicing classification as the independent variable (two levels: classified as voiceless or voiced). The effect of voicing was significant on all dependent variables (pooling all prosodic conditions: Closure Duration: $F(5, 20) = 5.989$, $p < .01$, Release duration: $F(5, 20) = 6.694$, $p < .01$, Relative Amplitude of Closure: $F(5, 20) = 4.365$, $p < .01$, Relative Amplitude of Release: $F(5, 20) = 4.433$, $p < .01$). Mean values for all measurements can be seen in Table 4.11.

The effect was in the anticipated direction for all variables except for the duration of the closure (see Table 4.11). The duration of the closure is longer in the voiceless than the voiced items, which is the opposite effect to the one anticipated. This is due to the fact that almost all voiced instances in this analysis come from BC0, the

lowest condition with the shortest possible duration, while all voiceless instances come from conditions BC1-BC5, which have been shown to have gradually longer durations. The duration of BC0 voiced closures was seen to overlap with BC4 and BC5 voiceless closures in the previous section too, and shows how it is possible to get a reversed effect when pooling data from across all boundary conditions.

In order to test whether this not anticipated effect is indeed due to BC0 having generally shorter durations, we tested the difference between stops classified as voiceless vs voiced only in BC0. In this condition the result should show the anticipated direction with voiceless closures being shorter than voiced ones. It was found that indeed, in BC0, the duration of the stops classified as voiced do have longer closures⁸, supporting the claim that the discrepancy found stems from the fact that all boundary conditions are grouped together, but voiced instances mainly occur in the lowest boundary condition.

The application of stop voicing, therefore, seems to take place categorically with clearly distinct resulting categories. Plotting the distribution of the values of ‘nasal+stop’ items classified as voiceless vs voiced against each other provides histograms with clearly distinct distributions (Figure 4.7, for example, shows the difference in the distribution between voiceless and voiced stops with regard to the amplitude of the release; there is substantially more data in the voiceless histogram since it plots all data from the experiment, and the voiced instances were fewer since they only occurred in BC0). A binary logistic regression analysis was performed to test which of the measured variables best predict the perceived voicing distinction. We saw in the first experiment that the acoustic characteristics of the closure accounted for voicing variation between phonologically voiceless and voiced stops best. For ‘nasal+stop’ sequences classified as voiceless vs voiced it was found that the acoustic characteristics of the closure were still significant in predicting the voicing variation, and also the relative amplitude of the release was added. Moreover, the duration of the closure was the least important of the three significant

⁸This was seen in the previous section, when we investigated whether there were acoustic differences between BC0 items classified as voiceless, voiced and no agreement. It was seen that the duration of the closure of BC0 items was not different for any of the three categories, with items classified as voiceless having a duration of 42ms, and those classified as voiced a duration of 43ms.

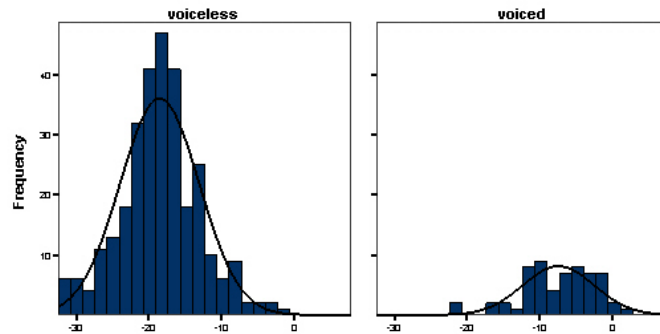


Figure 4.7: Histograms of the relative amplitude of the release with what was classified as voiceless on the left, and voiced on the right. The difference in mean values, and in the amount of voiced instances is apparent. This Histogram plots all measurements from all five boundary conditions.

variables. This is explained again on the basis that most voiced instances came from BC0, therefore the duration of the closure of voiced instances significantly overlapped with the longer closures of voiceless ones. It seems that, in the absence of much information from the closure, information from the release will be used for signalling voicing distinctions. Performing a discriminant analysis to see the success in predicting voicing contrasts using the predictor variables found to be significant in the regression model, 89.2% of the data were correctly classified. Using information from just the closure 86.7% of the data was correctly classified showing that the acoustic characteristics of the closure still manage to distinguish voiceless from voiced renditions, but not as well.

Findings thus far would suggest that the application of stop voicing is phonological, since its phonetic output seems to be categorically distinct between instances classified as voiceless vs voiced. However, it was mentioned in the previous section that there were also some instances that did not reach agreement within BC0. These were further analyzed statistically; we tested the statistical differences between the duration and amplitude of the closure and release of all stops in BC0 using the classifications as an independent variable. That is, we tested the acoustic differences between the items classified as voiceless, those as voiced, *and* those that didn't reach agreement, treating this last category as an extra level for the independent variable.

Four REPEATED MEASURE ANOVAS were performed, one for each variable. Results showed that the classification type was significant on the duration of the release ($F(2, 8) = 5.21, p < .05$), and the amplitude of both the closure and the release (Closure: $F(2, 8) = 5.23, p < .05$, Release: $F(2, 8) = 5.83, p < .05$). The duration of the closure was not statistically different across classification types (voiceless stops were 42ms, voiced 43ms and those not having reached agreement were 38ms). The ‘no agreement’ mean values for each variable were similar statistically to the voiceless ones (according to Bonferroni post-hoc tests), and different to voiced ones, suggesting that they had not undergone voicing, but there was something in their phonetic manifestation that made them resemble voiceless stops (Mean values of Duration of release: voiceless 20ms, voiced 15ms, no agreement 22ms, Mean values of Amplitude of Closure: voiceless -12dB, voiced -8dB, no agreement -13dB, Mean values of Amplitude of Release: voiceless -14dB, voiced -7dB, no agreement -12dB). From the current phonetic analyses the ‘no agreement’ cases, therefore, seem to resemble those of voiceless instances.

We also looked at the spectrogram and waveforms of the instances not having reached agreement once more, trying to identify the reason causing the disagreement across raters. For many of the wavefiles we could not identify the reason why the rendition would be ambiguous to some raters, since many of those seemed to be clearly voiceless perceptually, and others clearly voiced, at least to the experimenter. Many of other instances had a very pronounced pre-boundary nasal combined with a very short closure duration, but some of those items actually sounded voiceless. This might suggest an interesting interplay between the pre-boundary nasality, its articulatory character and interaction with the following closure, and its perception by the listeners. It might also correlate with whether the listener is actually expecting to hear a nasal when voicing takes place, or whether they are expecting it to be deleted, as Arvaniti (2004) proposed. All these issues, however, will need to be further addressed in an experiment testing the articulatory behaviour of the nasal.

To conclude, it has been shown, using both a perception experiment and a phonetic analysis of the classified items, that the distinction between instances classified as voiceless vs voiced from the ‘nasal+stop’ sequences is categorical, but that there

are instances that didn't reach agreement, and which suggest some type of coarticulation might be taking place. This is important since it shows that impressionistic analyses of the past have captured the realization of the phenomenon, but also that there might be more issues to be investigated before concluding regarding the nature of stop voicing. The main issue that needs further investigation is the issue of nasality, and how it might influence the production and perception of 'nasal+stop' sequences. Moreover, it is important to identify whether the reason for the difference in how many times voicing occurs across lexical items might be explained on the basis of articulatory analyses.

We now turn to the main investigation of this chapter, which is the effect of boundary strength on the application of stop voicing. Thus far all analyses have shown that the effect will be qualitative separating BC0 from the other conditions, in that in BC0 stop voicing is allowed. The next section will investigate the phonetic manifestation of 'nasal+voiceless stop' sequences to see whether there is a qualitative boundary strength effect apparent.

4.3.3.2 Qualitative or variable boundary strength effect?

The main goal of the present chapter has been to investigate the effect boundary strength exercises on the application of post-nasal stop voicing. The effect of boundary strength is investigated in the present section on all 'nasal+stop' instances, on all measurements. Two sets of hypotheses are formed; first, how the phonetic characteristics of underlyingly voiceless sounds will change if stop voicing occurs qualitatively, i.e. in the lowest of the domains and not in the higher domains as predicted by the perception experiment, and second, how the same phonetic characteristics will change if stop voicing occurs variably, that is if more instances of stop voicing occur in lower domains and less instances in higher domains.

Figure 4.8 shows the hypotheses of how each variable is expected to be influenced if stop voicing mirrors a qualitative or variable boundary strength effect. The Figure is broken down in two panels by a vertical black line. The left panel is the type of effects found in the production experiment of voiceless and voiced stops to be exercised on voiceless stops on each measured variable (can be compared to the trends shown in Figure 4.5). The right hand panel shows how this effect is expected to vary

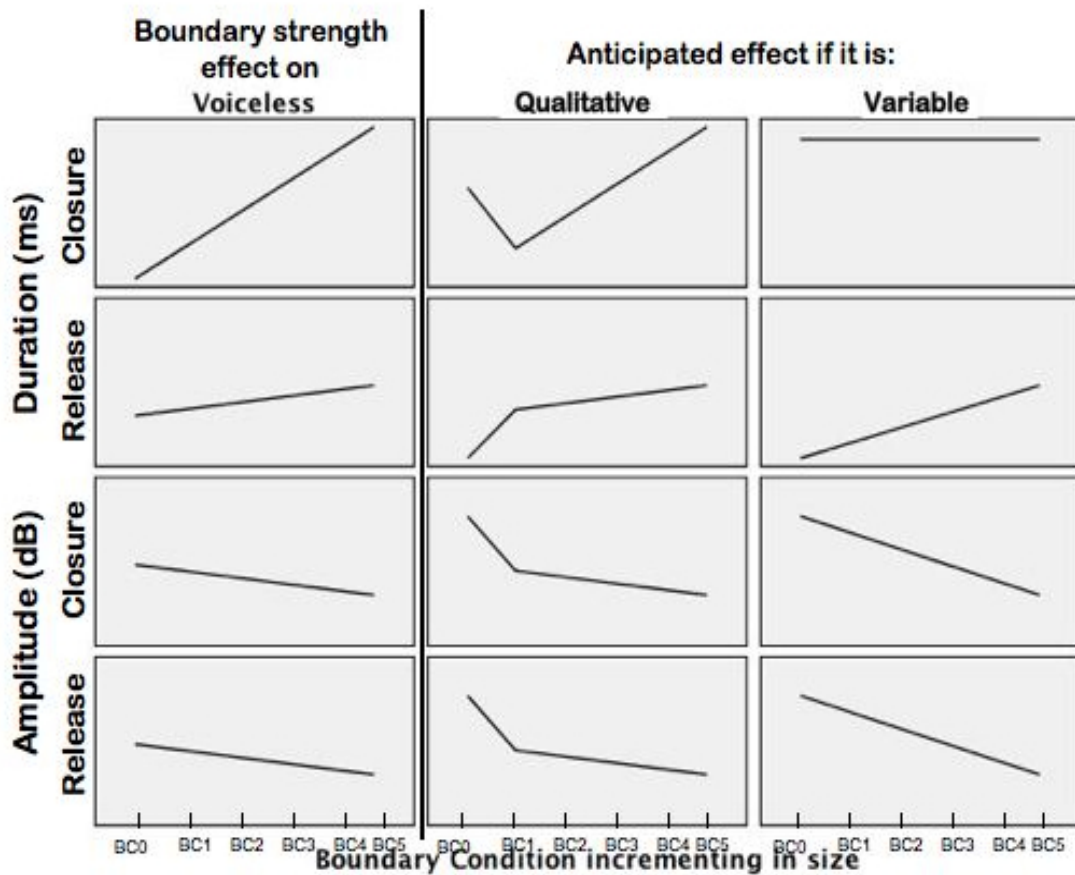


Figure 4.8: Hypotheses of how boundary strength will effect the phonetic manifestation of each variable if stop voicing occurs qualitatively (second column), or variably (third column). The first column (separated from the others by a black vertical line) shows the effect found to be exercised by boundary strength on voiceless stops. Based on that expectations are formed. The x axis shows boundary conditions incrementing in size, and the y axis shows Duration (ms for the first two rows) and Amplitude (dB for the last two rows) measurements. Figure is explained further in text.

if, on top of it there are also qualitative (left column in right hand panel) or variable (right hand column in right hand panel) boundary strength effects on the variables. By qualitative boundary strength effects it is meant that stop voicing occurs in the lowest boundary condition but not in the others, supporting the existence of a qualitatively distinct type of prosodic domain. By variable boundary strength effect it is meant that there are gradually fewer instances of stop voicing higher in the structure. The rows in this Figure represent the four measured variables; the duration and amplitude of the closure and release. In the behaviour of the lines

	Prosodic boundary effect						
	Prosodic Condition						Post-Hoc
	BC0	BC1	BC2	BC3	BC4	BC5	
Closure Duration(ms)	42	34	45	41	45	51	1<0=3245,3<5
Release Duration(ms)	17	24	23	24	23	25	0<12345
Closure Amplitude(dB)	-9	-15	-14	-15	-14	-12	0<12345
Release Amplitude(dB)	-10	-16	-18	-19	-20	-22	0<12345,1<5

Table 4.12: Mean values for all dependent variables in all prosodic conditions. The last column shows whether the effect of prosodic boundary strength was significant for each variable.

in the first column (the phonologically voiceless renditions) it is shown that there is a rather marked effect on the duration of the closure from boundary strength, while the other variables are not that influenced. In the middle column we can see that, if there is a qualitative difference in the number of times stop voicing takes place, the duration of the closure of the lowest domain will be markedly longer, and the duration of the release markedly shorter; this means that the mean duration of the closure and release will resemble more that of voiced stops (longer closures and shorter releases). The amplitude of both the closure and the burst will be markedly higher, since voiced stops have higher amplitude. If there is a variable effect (last column), the duration of the closure should look very similar across conditions if there are gradually fewer instances of stop voicing; more voiced instances at BC0 will make the duration of the closure very long, while the least instances in BC5 will preserve the initial duration BC5 has when it comprises of solely voiceless renditions. The duration of the release will be gradiently shorter, while the amplitude of the closure and release will be gradiently lower, signifying less voicing in higher domains, therefore fewer instances of voicing. The fact that the trend for shorter duration and less amplitude in those last three variables is more pronounced than in the first and second columns is derived exactly from the fact that there is a gradual increase in the number of voiced instances across domains in this last column, which represents the variable boundary strength effect.

The effect of the strength of the prosodic boundary on each of the dependent variables was tested using four REPEATED MEASURE ANOVAs, one for each measured

variable. Table 4.12 shows the mean values for each variable in all prosodic conditions. The effect of prosodic boundary strength was significant on all measured variables: Duration of Closure: $F(5,424) = 9.5$, $p < .001$, Duration of Release: $F(5,481) = 9.8$, $p < .001$, Relative Amplitude of Closure: $F(5,477) = 14.2$, $p < .001$, Relative Amplitude of Release: $F(5,419) = 37.3$, $p < .001$. Table 4.12 shows the planned Post-Hoc Bonferroni comparisons. It is shown that the behaviour of BC0 in all measured variables confirms the expectation that post-nasal stop voicing is influenced in a qualitative manner, giving rise to two distinct domains; the lowest vs all the others, and confirming the result from the perception experiment. The duration of BC0 is longer than that of BC1, and not different to BC2-BC5, showing how for BC0 there is an effect making the duration of the closure longer. Also, across conditions BC1 and BC5 there is gradual lengthening of the duration of the closure, as anticipated from a potential post-boundary effect. The duration of the release again confirms the qualitative effect hypothesis, since it is markedly shorter than all other conditions. Both amplitude measurements for the closure and the release also confirm the expectations, with BC0 assuming the highest value (therefore has the highest amplitude) of all other domains. Finally, it is clear that, if any post-boundary effect takes place, it is mainly confined to the closure of the stop, as anticipated from all findings in the previous sections.

4.3.3.3 Deletion of pre-boundary nasal

In the Introduction to this chapter it was mentioned that the behaviour of the preboundary nasal might be a significant cue as to whether stop voicing occurs or not. However, given the lack of nasal and oral airflow measurements for this experiment, it was not possible to address this issue. Therefore, I performed an impressionistic analysis of where nasal deletion might have taken place, as an initial means of understanding the behaviour of the nasal. As this measurement relies upon my classifications, and is therefore rather arbitrary, it is regarded as additional evidence for the categoriality of the post-nasal stop voicing application - not as proof of voicing application, and provides tentative conclusions for further research.

Table 4.13 shows the percentage of nasal instances transcribed as having deleted the pre-boundary nasal in the 'nasal+stop' sequences, broken down by prosodic

Deleted?	Boundary Condition					
	BC0	BC1	BC2	BC3	BC4	BC5
<i>Yes</i>	37%	2%	3%	2%	1%	1%
<i>No</i>	63%	98%	97%	98%	99%	99%

Table 4.13: Percentage of nasals being classified as deleted or not across boundary conditions.

condition. There is a difference of 35% between BC0 and BC1, with 37% instances indicated as having undergone nasal deletion in BC0, and only 2% being marked for BC1 and the following conditions. It is clear that deletion of the pre-boundary nasal is mostly present in the lowest of the experimental conditions, suggesting a categorical difference, as the one expected by Arvaniti & Joseph (2004).

4.4 Conclusions

Studies on prosodic constituency have used sandhi processes to support the existence of labelled prosodic domains. The blocking of the application of a sandhi process across two domains is viewed as support for the existence of the purported domain, giving rise to a binary distinction between the domain within which the phenomenon is allowed to occur, and the domains that block it.

Within this view, Nespor & Vogel (1986) used post-nasal stop voicing from Modern Greek to postulate the existence of specific domain types within their structure. However, their analysis suffered from two assumptions. First, they performed impressionistic analyses and assumed that the phenomenon is a sandhi process, when it could be the result of higher or lesser degrees of coarticulation. The present chapter aimed at investigating the acoustic realisation of post-nasal stop voicing, and to indirectly discuss whether its application in Modern Greek is a phonological process, before testing the effect boundary strength exercises on it. Second, in Nespor and Vogel's account, all boundary strength effects on sandhi-type phenomena were anticipated to be qualitative, while in reality the effect of boundary strength on phenomena can be variable. Therefore, the present chapter aimed at testing

whether stop voicing is qualitatively influenced by boundary strength, and subsequently whether Nespor & Vogel's (1986) analysis holds, thus providing evidence that qualitative differences across domains can be identified. Finally, Nespor and Vogel investigated stop voicing in a limited number of prosodic conditions, since they only used it for the postulation of the phonological word and clitic group. The present analysis, however, investigates the behaviour of the phenomenon under a variety of boundary conditions, therefore testing whether qualitative distinctions between domains arise, as a qualitative view would predict, or whether variable differences emerge which only appeared to be qualitative in Nespor & Vogel's (1986) analysis given the very limited number of conditions investigated.

Regarding the general question of whether the application of stop voicing appears to be qualitatively or variably influenced by boundary strength, the present experiment provided solid evidence in support of Nespor & Vogel's (1986) claims that it is affected by boundary strength in a qualitative manner. A perception experiment showed that 62% of 'nasal+stop' sequences were classified as voiced in BC0 (the lowest condition investigated), while in conditions BC1 and BC5 there was a radical drop to 0-9% of the instances being voiced within each condition. This showed that the application of stop voicing was optional within BC0, and blocked across.

The data in BC0 was broken down by lexical item; Nespor and Vogel's analysis of stop voicing in Modern Greek had suggested that they expected a distinction between the temporal particle ['otan] and the negation particle [ðen] in Modern Greek, a finding that again was supported by the present data, in the sense that voicing occurred in all instances containing a negation particle, and in only a very small fraction of the instances containing the temporal particle. Nespor and Vogel's analysis predicts that stop voicing is obligatory within a phonological word, optional within a Clitic Group, and blocked across Clitic Groups. Therefore, they predict that ['otan] forms a clitic group on its own, since it is blocked, while [ðen] forms a phonological word joining with the following verb to form a Clitic group, therefore voicing is *optionally* allowed. However, voicing was found to always occur for this particle, as would be expected within phonological words. To add to the picture, the other temporal particle [min] only voiced approximately 50% of the times, fitting Nespor and Vogel's optionality proposal.

It is, therefore, clear that while the present experiment has provided support for the postulation that there are qualitative boundary strength effects influencing the application of sandhi processes, the way this grouping takes place is far more complicated than just accounting for how phonology maps on to syntactic constituency. While this chapter has managed to answer the main research question, it has also raised further questions, such as the need for an explanation for the differences found across lexical items. Irrespectively of the above, however, the present chapter offers clear perceptual and production data in support of views entailing labelled distinct constituents.

The phonetic implementation of post-nasal stop voicing was also investigated in this chapter, and it was shown that the phonetic differences between instances of ‘nasal+stop’ classified as voiceless vs voiced fitted neatly two distinct categories resembling those of phonologically voiceless and voiced stops, supporting the view that it is a phonological sandhi process. However, this is not a conclusive investigation, given that articulatory measurements were not taken, and given that there were instances for which raters did not agree whether voicing had taken place or not, and for which it was not clear whether articulatorily there is gradience or not. Moreover, it might be the case that there are quantal effects on the articulation side of the production of ‘nasal+stop’ sequences, i.e. very small changes in the articulatory continuum could give rise to categorical differences acoustically and perceptually across instances classified as voiceless and voiced. Still, however, these can be viewed as categorical cues to prosodic constituency. Finally, there seems to be some linguistic notion regulating where stop voicing occurs, given the differences across lexical items found (e.g. how often stop voicing occurred across [an] ‘if’ vs [ðen] ‘not’ particles). However, those are issues that require further experimental investigation.

In the process of answering the main research question of how boundary strength affects stop voicing application, the present chapter addressed issues that had not been examined experimentally in the past for Modern Greek. First, acoustic distinctions between voiceless and voiced stops were addressed. It was found that all measured variables were significantly marking the phonological contrast of voicing;

the duration of the closure was longer for voiced than voiceless stops, the duration of the release was longer for voiceless than voiced stops, and the amplitude of the closure and the release was higher for voiced than voiceless stops. Moreover, it was found that the duration and amplitude of the closure are more important in signalling the phonological contrast of voicing, while the relative amplitude of the release also became significant in signalling the contrast of voicing between ‘nasal+stop’ instances having undergone voicing or not, given that the duration of the closure in those latter instances was somewhat ambiguous.

These findings, along with the relatively small effects of boundary strength on post-boundary voiceless and voiced stops, were taken to support results from Chapter 3, which reported that the effect of boundary strength in post-boundary segments in Modern Greek is not significant. While it gives rise to some lengthening, it is mainly localised in the immediate vicinity to the boundary, in this case the closure of the stops. Moreover, when it did take place, its effect was not so strong as to interfere with voiceless vs voiced distinctions, therefore in almost all cases the two categories are kept distinct.

Finally, it is important to note that in the present chapter we have only investigated the phonetic implementation of the post-boundary stops, while the pre-boundary nasal has not been investigated. Similarly, measurements of nasality within the closure of the stops have not been taken. The issue of prenasalisation has received a lot of attention in the Greek literature, and it could be a further cue to voicing distinctions in ‘nasal+stop’ sequences, as our tentative analysis of where nasal deletion might have occurred showed in the present chapter. To sum up, not much consensus exists as to whether voiced stops in Modern Greek should be considered as being preceded by a full homorganic nasal (e.g. [mb]), or whether they are prenasalised segments (e.g. [ᵐb]). Recent work has discussed that voiced stops are not always preceded by a nasal (Pagoni-Tetlow 1993, Arvaniti & Joseph 2004). Rather, it has been proposed that either the nasality is connected with the stop following it, or in new generations of native Greek speakers the nasality is no longer there (Arvaniti & Joseph 1999, Arvaniti & Joseph 2004). Nasality has also been found to vary in extent, sometimes lasting almost as long as the oral closure but at others being present only for a brief period of time at closure beginning. For a full discussion

of the nature of prenasalisation in Modern Greek, the interested reader is referred to Arvaniti & Joseph (2004), Arvaniti (2007), Malikouti-Drachman (2001), Pagoni-Tetlow (1993), among others. The issue of nasality might assume an important role when looking at how stop voicing takes place in BC0, since it might be the case that nasality becomes very important in understanding whether it is a sandhi or a coarticulatory phenomenon. The nasal could be completely deleted, supporting the existence of a phonological phenomenon, and nasality could also be influenced qualitatively or variably by boundary strength, adding to the conclusions drawn in this chapter.

CHAPTER 5

Vowel Hiatus resolution

This chapter is closely linked to the previous one, since its subject matter is the effect boundary strength exercises on a sandhi-type process used in labelled views for prosodic domain identification, namely vowel hiatus resolution. Therefore, the motivation behind its investigation is very similar to that of the previous chapter, with one difference: while stop voicing had not been analysed instrumentally, and was found in Chapter 4 to behave like a phonological process (at least in terms of its application within the prosodic structure), vowel hiatus resolution has been the matter of analysis in several recent studies which have reported that its resolution is the result of coarticulation. Therefore, the present chapter sets out to test the effect boundary strength exercises on a process which is most likely the result of coarticulation.

A perception and production experiment are used, similarly to Chapter 4. The preliminary perception experiment investigates how vowel hiatus resolution is perceived by transcribers, and tests the reliability of impressionistic analyses for this phenomenon. The production experiment is used to provide further evidence for the phonetically gradient nature of the phenomenon, and then for investigating the effect boundary strength exercises on formant trajectories and the duration of the vowel pairs; these are the two types of variables used for measuring the amount of coarticulation across boundary conditions.

5.1 Introduction

5.1.1 Definition and nature

Hiatus is created when two vowels are immediately adjacent. For example, in (5.1) the vowels [ao] (in bold) are in hiatus (example from Modern Greek):

(5.1) 'omorfa **o**'nomata_{NOM-PL} (beautiful names)

In the present chapter only *external* vowel hiatus will be investigated, that is two vowels separated by a word boundary.

Hiatus has been said to be avoided cross-linguistically (Casali 1997), which means that languages often have *resolution strategies* for resolving hiatus. In traditional phonological literature it has been said to be resolved with a variety of phonological phenomena, like vowel assimilation or vowel deletion. For example, vowel hiatus resolution in Igbo has been reported to be resolved using vowel assimilation (literature cited in Zsiga 1997), as in example (5.2).

(5.2)

a. V1V2 → V1V1 (if V2 is assimilated to V1)

b. V1V2 → V2V2 (if V1 is assimilated to V2)

nwoke a → nwoka a

man DEF 'this man' (example taken from Zsiga 1997, pp. 237)

External vowel hiatus in Modern Greek has traditionally been analysed as being resolved using vowel deletion:

(5.3)

a. V1V2 → V1 (if V2 deletion has taken place)

/to 'exo/ → ['toxo]

'(I) have it'

b. V1V2 → V2 (if V1 deletion has taken place)

/to 'afisa/ → ['tafisa]

'(I) left it'

Phonological research in Modern Greek has proposed that vowel hiatus is resolved using deletion, usually of the first vowel of the vowel pair (V1) (Hatzidakis 1905,

Kaisse 1985, Condoravdi 1990, Nespor & Vogel 1986, Malikouti-Drachman & Drachman 1992). Resolution via deletion is believed to take place optionally, and mainly in fast speech. Moreover, traditional analyses of vowel deletion in Modern Greek have proposed the existence of a *sonority hierarchy* in Modern Greek according to which some vowels are *stronger* than others in the sense that they are less likely to be deleted. One version of the hierarchy proposes that moving from strongest to weakest the hierarchy is [o a u i e] (Hatzidakis 1905, Kaisse 1985), while Malikouti-Drachman & Drachman (1992, pp. 201) propose [a o u i e]; that is, the difference lies in the relative position of [a] and [o]. Such disagreements highlight how traditional phonological analyses have not managed to capture the way vowel hiatus is resolved in Modern Greek.

Thus, traditionally, research dealing with resolution of non-identical vowel hiatus assumed that all resolution strategies available to speakers are of phonological nature. This assumption relies on impressionistic analyses of the phenomenon, without instrumental or articulatory support. Recent research, however, investigating acoustic and articulatory characteristics of sandhi-type processes in several languages is in direct opposition to such analyses, showing that the phonetic output of processes traditionally classified as assimilation or deletion can be one of the possible outcomes of a gradual phonetic process than of a discrete change from one type of sound to the other (Nolan 1992, Nolan et al. 1996, Ellis & Hardcastle 2002). In particular regarding vowel hiatus resolution, it has been proposed that the process can be the output of coarticulation and can assume gradient, continuously variable phonetic values, therefore not necessarily involving discrete phonological changes from one sound to the other.

One of the best-known studies investigating the continuous nature of vowel hiatus resolution is that by Zsiga (1997). She concurrently investigated two vowel processes, namely vowel harmony and vowel assimilation, in Igbo. Vowel harmony was reported to be best represented in autosegmental terms involving a categorical rule, while vowel assimilation was found to be a phonetically gradient process best represented in purely articulatory terms. She investigated phonetically a number of vowel hiatus pairs, and found that the most common resolution strategy in Igbo is partial assimilation, that is the formant values of the vowel pair investigated had

values between the baselines; for example, in a vowel pair [ao] the formants at the beginning of the vowel pair would lie inbetween the baseline formant values of [a] and [o]. This contradicted previous analyses of the phenomenon for that language, which proposed that when hiatus occurred V1 assimilated completely to V2. Her durational data also provided evidence for partial assimilation, not deletion, since the stretch of V1V2 vowel pairs undergoing assimilation was similar to the vowel pairs that didn't, and not shorter as would be expected if deletion had taken place. Based on those two processes (one being phonological and the other phonetic) she proposed a link between the autosegmental and gestural representations of the two processes, where both representations together can account for the discreteness of phonological rules and the articulatory overlapping of gestures. Importantly for the present discussion, she provided solid evidence that vowel hiatus resolution is not necessarily achieved by means of categorical phonological processes.

With respect to vowel hiatus resolution in Modern Greek, there have been recent studies reporting that hiatus resolution is best regarded as a phenomenon with varying phonetic output. Pelekanou & Arvaniti (2001), using a corpus of naturally occurring speech, supported the view that the vowel hiatus resolution is phonetically gradient, with a number of possible outcomes ranging (using their terminology) from deletion to coalescence, centralisation, etc. This result was corroborated by Baltazani (2006), who, in a carefully controlled experiment with laboratory speech also reported on gradient phonetic output of vowel hiatus resolution, and supported the impression that partial assimilation is the most common phonetic output of vowel hiatus resolution in Modern Greek. She had grouped the potential phonetic outcomes in four categories: deletion, assimilation, merging and no assimilation, and she classified each item manually into one of these categories (however, her criteria for classification are not clear). Using this classification technique for representing the distribution of possible phonetic outputs, she showed that all different outcomes are possible within her data, countering the traditional impression that V1 deletion is the most common strategy. Finally, in her experiment, the duration of the V1V2 stretches correlated with formant values, in the sense that longer durations coincided with more extreme formant values, suggesting that the phenomenon might indeed

be best described as a coarticulatory phenomenon of gestural overlap (a conclusion shared with Pelekanou & Arvaniti, 2001).

Recent studies on vowel hiatus resolution in Modern Greek, therefore, provide evidence for the inadequacy of impressionistic transcriptions to account for the true nature of vowel hiatus resolution. Moreover, they have shown how traditional impressionistic analyses have not managed to capture some of the findings acoustic analyses report on. For example, Baltazani's (2006), Pelekanou & Arvaniti's (2001) and Kainada's (2007) analysis of several vowel pairs showed that both V1 and V2 deletion are possible in Modern Greek (with V1 deletion being a bit more common in Baltazani's data). Moreover, they have shown that neither of the sonority hierarchies proposed by traditional research is replicated in acoustic analyses of vowel hiatus resolution. Pelekanou & Arvaniti (2001), Baltazani (2006) and Kainada (2007) investigated [ao] and [oa] vowel pairs, and reported that both [a] and [o] can be deleted in either vowel pairs¹, thus again providing evidence against impressionistic analyses of vowel hiatus resolution.

In conclusion, vowel hiatus resolution in Modern Greek has been shown to be a coarticulatory phenomenon. When vowel hiatus is resolved, it can range on a continuum of possible outputs. The present experiment will first investigate how vowel hiatus is resolved in the present production experiment to confirm that hiatus resolution does seem to be the result of coarticulation. This will be achieved by testing how formant values of non identical vowel pairs compare to baseline vowel pairs with identical vowels, testing whether there is variability in the outputs, or whether the produced non identical vowel pairs resemble the identical ones, suggesting deletion/assimilation is the resolution strategy used. The goal is to confirm that vowel hiatus is not a phonological sandhi phenomenon involving a categorical effect of, say, deletion, and therefore investigate boundary strength effects on that phenomenon. The following section focuses on how vowel hiatus resolution has been treated in

¹However, the vowels involved in hiatus were found to be a crucial factor for the type of resolution strategy in Baltazani (2006), who reported that [ao] sequences have a high rate of V1 deletion (39% of occurrences), while [ua] sequences hardly ever underwent V1 deletion. Baltazani was not able to propose an explanation for this finding, and it remains to be seen what is the regulating factor for these differences (prosodic, articulatory, syntactic etc).

the past with respect to its use for prosodic domain identification within labelled views, and then in more recent studies.

5.1.2 Vowel hiatus resolution and prosodic domain identification

Vowel hiatus resolution is a process extensively referred to in theories of prosodic phonology supporting the qualitative view on prosodic structure. Nespor (1987) used vowel deletion to support the existence of the Clitic Group. For example, based on deletion of V1 in example (5.4) they propose that the two words belong to a Clitic group, since vowel deletion is allowed.

(5.4) $\tau\alpha \acute{\epsilon}\chi\omega$ [ta 'exo] \rightarrow [t'axo] '(I) have them'

However, as Arvaniti (2007) points out, their analysis does not capture the lack of deletion, for example, between $\tau\alpha \epsilon\delta\acute{\alpha}\varphi\iota\alpha$ [ta e'ðafia] \rightarrow *[t'a'ðafia] 'the passages' (see Arvaniti 2007, pp. 70ff, for a discussion of cases not captured by such analyses). In a similar analysis to that of Nespor and Vogel, Condoravdi (1990) used deletion rules of non-identical vowels to propose the existence of a further prosodic domain between the Prosodic Word and the Phonological Phrase, the *Minimal Phrase* (she did not have a Clitic Group in her analysis). However, as Arvaniti (2007) discusses, some of the proposed sequences she discusses are impossible for native speakers of Greek, going against their intuitions of whether deletion can occur or not.

Therefore, traditional analyses of the phenomenon of vowel hiatus resolution within labelled views have not managed to present a proposal covering all possible factors influencing its application. Moreover, their reliance on impressionistic data derived solely from the experimenters' classifications create problems for the proposed analyses in that they often do not represent native speakers' intuitions.

The way the phenomenon has been used in labelled views to provide support for distinct domain types has further been questioned by subsequent instrumental research. Given the finding that the nature of the phenomenon is coarticulatory, the question that arose was whether the effect of boundary strength on its application might itself be variable (instead of qualitative, as proposed by past analyses within labelled views). Recent studies on the phenomenon tested instrumentally the effect boundary strength exercises on it (Pelekanou & Arvaniti 2001, Baltazani 2006,

Kainada 2007). However, they have provided mixed results with respect to how boundary strength influences vowel hiatus resolution across prosodic conditions. In some studies it is claimed that there is a qualitative effect, but they disagree on exactly the domain across which blocking occurs, while others support the view that the effect is variable across domains with more coarticulation lower in the structure and no categorical effects. Pelekanou & Arvaniti (2001) reported that vowel hiatus resolution can assume varying outputs, but these are blocked across what they name intermediate phrases, supporting the view that the effect is qualitative. Baltazani's (2006) analysis of vowel hiatus (also mentioned in the previous section for her insights into the nature of vowel hiatus resolution), on the other hand, which specifically set out to investigate the effect of boundary strength on the phenomenon, reported that the effect is blocked across what she annotated as intonational phrases, supporting again the view for a qualitative effect but at a different level. While Baltazani mentions that there seems to be an IP blocking effect, she does not clearly conclude if it is a qualitative effect or not. The IP effect she reports, however, might be an artifact of the methodology she used; formants were measured from the whole V1V2 stretch. She reports that, if there was a silence between the two vowels, the silence would be chopped out and the two vowels combined for the purposes of measurements. However, this suggests that any coarticulatory patterns will be lost in the presence of silence. Furthermore, given that pauses are more likely to occur in higher domains, it is expected that the IP condition (the highest condition involving a boundary across two sentences separated by a full-stop) is very likely to give rise to a categorical difference to the other conditions. However, she does not report on how many instances with silence came up in her data, nor on their distribution across conditions. Therefore, it is inconclusive whether a real qualitative effect is identified or not.

Baltazani's (2006) experiment is very similar to the one presented in this chapter, with experimental conditions very closely related to ours. She examined eight vowel pairs separated by an orthographic word boundary. The pre-boundary word for some of the vowel pairs was an Adjective and the post-boundary word a Noun, exactly like the design of the current study. The items were placed under three prosodic conditions, namely PrWd, ip, and IP, based on Arvaniti & Baltazani's

(2005) analysis of GrToBI. PrWd marks a distinction between e.g. an Adjective and a Noun (comparable to our BC1), ip marks a subordinate clause next to a main clause (comparable to our BC3), and finally IP marks a separation between two main clauses separated by a full-stop (comparable to our BC5). Example (5.5) shows the prosodic contexts in which the items were embedded².

(5.5) **PrWd:** *Tha sou etoimaso ena nostimo*]Wd *uzaki me meze.*

will for-you prepare-1s one tasty ouzo with tidbits

‘I’ll prepare tasty ouzo with tidbits’

ip: *An exei meze nostimo,*]ip *uzaki pinun oloi.*

If it has tidbits tasty ouzo-dim drink-3p all

‘If the tidbits are tasty, everyone drinks ouzo’

IP: *Tha fao mezedaki an einai nostimo.*]IP *Uzaki den pino.*

Will eat-1s tidbits-dim if it is tasty ouzo-dim not drink-1s

‘I’ll have tidbits if they are tasty. I won’t drink ouzo.’

Her findings can, therefore, be directly compared to the ones from the present design, since three of her conditions are identical to three of ours (as was also mentioned in Chapter 2). The present study adds to the number of boundary conditions used, and will hopefully unveil more boundary strength gradations in support of any types of effects found.

Pelekanou & Arvaniti’s (2001) and Baltazani’s (2006) studies, therefore, seem to agree on the fact that boundary strength does exercise a qualitative effect on vowel hiatus resolution, but disagree on the level distinguished. This mismatch can come from many reasons: first, one study uses laboratory speech and the other naturally occurring speech, and second, the vowel pairs used across studies are not the same. Baltazani specifically reported that the vowels comprising the vowel pair might be a significant factor in the way resolution takes place. The number of items both

²One issue to note, is that the in Baltazani’s experiment the part-of-speech of the words preceding and following the word boundary was not the same across vowel pairs; for example, [ao] pairs were made up of *Adjective+Noun* constructions, while [oa] pairs were made up of *Verb+Object* constructions. If syntactic structure has any influence on the prosodic grouping of the utterances involved, the two conditions are not comparable. Such a syntactic influence is not expected within the intonational approach, but the Syntactic Approach and the direct syntax-phonology mapping approaches might predict different prosodic structures.

studies analysed was small, divided into a number of conditions, and the part-of-speech of the words in hiatus was not controlled for.

One final study testing the effect of boundary strength on vowel hiatus resolution is that of Tserdanelis (2004). He used identical vowel pairs (an important difference to the previously mentioned studies, since hiatus in those can also be resolved by vowel degemination) and tested whether speakers' productions change across syntactically ambiguous sentences with an adjective scope ambiguity (example 5.6 from Tserdanelis 2004, pp. 53). His hypothesis was that intonation patterns and segmental sandhi will disambiguate the renditions and support the existence of two distinct prosodic domains.

(5.6) Ἦρθαν νέοι υπάλληλοι και ιδιοκτήτες στη συγκέντρωση
/i'ρθan 'nei i'palili ce iðio'ktites sti si'gedrosi/

(They) came new employees and employers to the reception

Meaning a. *employers and new employees came to the reception*

Meaning b. *both employers and employees were new*

However, his results did not support the hypothesis that vowel hiatus resolution strategies would differ between his syntactically ambiguous conditions. Moreover, speakers did not use intonation to disambiguate the syntactic ambiguity, a fact that was commented by Arvaniti (2007) suggesting that the syntactic ambiguity might be too difficult for native speakers to understand and produce. Therefore, while Tserdanelis concluded that the effect of boundary strength is not categorical on vowel degemination (as Nespor & Vogel 1986, would predict), it is not clear whether the finding is due to a lack of a qualitative effect, or to the speakers not being able to produce both interpretations of the ambiguous materials.

Thus, the present study aims at providing further evidence for whether the nature of vowel hiatus resolution in Modern Greek is truly a sandhi-type process or a coarticulatory phenomenon, and, importantly, to address the issue of whether the effect of boundary strength on the phenomenon is qualitative, as proposed by traditional analyses and by Pelekanou & Arvaniti (2001), or variable, as seems to be the case in Baltazani's (2006) data. The answer to this question will also provide insights to the general question of whether prosodic structure consists of labelled domains,

and will be compared to the other three phenomena investigated in this thesis at the Discussion chapter.

5.2 Goals and Overview of Chapter

The goals of this chapter are:

1. to test whether the resolution of vowel hiatus in Modern Greek entails discrete changes from one sound to the other, supporting views analysing it as a sandhi-type phenomenon, or gradient continuous phonetic output ranging in a number of possible outcomes, thus supporting its analysis as a coarticulatory phenomenon. The answer to this question is important in whether the phenomenon is treated as phonological or coarticulatory when identifying boundary strength effects on it. In the case of a sandhi phenomenon, the phonetic output is expected to fall into two distinct categories, depending on whether the phenomenon takes place or not. Therefore, we are testing the distribution of how many times the phenomenon takes place across boundary conditions, or we are looking for a categorical change in the distribution of phonetic values across boundary conditions showing where the phenomenon is allowed and where it is blocked. If the phenomenon is coarticulatory, on the other hand, the interest shifts to investigating whether the amount of coarticulation changes across boundary conditions. In the case of vowel hiatus, a categorical boundary strength effect, for example, would allow variable degrees of resolution in some domains and completely block it across others. A variable boundary strength effect would change the amount of coarticulation across conditions expecting more coarticulation (in this case measured in the position of formants at the beginning and end of the vowel trajectories) lower in the prosodic structure, and less coarticulation (more ‘typical’ positions of formants) higher in the structure.
2. Having identified the nature of the phenomenon, to investigate the effect of boundary strength on the phenomenon with the goal of testing whether it offers support for the existence of labelled domains. In particular, it is aimed at testing whether a variety of boundary strength conditions can give rise to a qualitative effect, like either of the one reported in Pelekanou & Arvaniti

(2001) or Baltazani (2006), or whether it will show how by testing a variety of boundary strength contexts one can identify a variable boundary strength effect. Moreover, the boundary strength effect analysis is set to test whether past impressionistic analyses à la Nespor & Vogel or Condoravdi were right in proposing distinct labelled domains on the basis of this phenomenon.

The two goals are addressed using two experiments; a preliminary perception experiment, and a production experiment. The perception experiment tests how much transcribers agree in classifying vowel pairs in hiatus into pre-defined categories, as impressionistic traditional analyses have done. These results also provide a first insight into boundary strength effects on how many times vowel pairs were classified in each category. Following that, we present the production experiment, which is divided in two parts; the first one looks at the phonetic output of the vowel pairs under investigation compared to baselines, testing whether the formant trajectories of vowels in hiatus fall into neatly distinct categories resembling those of the baselines, or whether there are variable outputs across items ranging in a continuum of possible values. This will address the question whether the resolution of hiatus is categorical or gradient phonetically. The second performs statistical analyses on the raw formant and duration data from the [ao] and [oa] vowel pairs to test whether there is a significant boundary strength effect, and in what way it takes place. This analysis will conclude as to whether coarticulation is blocked across some levels, supporting labelled views, or whether it is affected variably.

5.3 Methodology

5.3.1 Production experiment

5.3.1.1 Materials

In order to create the appropriate vowel hiatus context, two non-identical consecutive vowels are needed, one in pre- and one in post-boundary position. The preboundary word, as for all phenomena investigated in this thesis, is always an ADJECTIVE and the post-boundary word is always a NOUN. Two different pairs of vowels were used: [ao] and [oa] (see example (5.7)).

(5.7) [oa]: [a'notato a'poktima]_{supreme asset}

[ao]: [a'ðinata o'moloɣa]_{weak bonds}

The reason for choosing these particular vowel pairs was twofold: first, the morphological endings of [a] and [o] are very common in Greek, making it easy to construct materials; second, the formant trajectories for F1 and F2 for both [ao] and [oa] are moving in the same direction, making it easier for the researcher to identify them during segmentation, and between vowel pairs they are moving in opposite directions making any comparisons between them very clear. Two theoretical questions are also addressed by using those two vowel pairs; first, as it was noted in the introduction of this chapter, there has been a lot of debate regarding which vowel is *stronger* between [a] and [o] in Modern Greek, and using the two in interchangeable positions will allow us to investigate this question; second, having the same vowels alternate positions allows to examine whether the position of the vowel in the vowel hiatus is important. For example, if there are differences in the duration between vowel pairs, this could be an indication of influence from the position on the resolution of hiatus (a question that specifically relates to the expectations set out by Casali (1997) regarding those two vowel pairs, who proposed that V1 has some special stronger status cross-linguistically).

The last vowel of pre-boundary word is referred to as **V1**, the initial vowel of post-boundary word as **V2**, and the vowel pair as **V1V2**. Segmentally, before V1 there was always a voiceless dental plosive (/t/), while V2 was always followed by either a voiceless bilabial plosive (/p/) or a voiced bilabial nasal (/m/). The same segments were used so as to avoid the confounding factor of different formant transitions into and out of the V1V2 sequences between materials. Table 5.1 shows the six carrier words created for each vowel pair.

Finally, materials for baseline conditions were created using the vowel pairs [aa] and [oo]. These could not be directly compared in terms of durations and formant trajectories, given that they might be undergoing different resolution strategies (e.g. more frequent deletion of V1 or V2). However, they do provide the baseline formant values for [a] and [o] from each speaker. Again six carrier words were created for each vowel pair (Table 5.1).

/ao/	1	[a'ðinata o'moloɣa] weak bonds	/oa/	1	[i'pertato a'pofθeyma] supreme saying
	2	[ka'totata o'moloɣa] lowest bonds		2	[a'notato a'poktima] supreme asset
	3	[a'nakata o'moloɣa] mixed bonds		3	[en'ðekato a'postolo] eleventh apostle
	4	[θe'orata o'mometra] huge omometers		4	[a'ðinato a'poɣono] weak descendant
	5	[θe'orata o'palia] huge opals		5	[a'ðinato a'moni] weak anvil
	6	[a'ðinata opti'ka] weak specs		6	[p'rosfato a'molutro] recent sand-bath
/aa/	1	[a'notata apo'ktimata] supreme assets	/oo/	1	[a'ðinato o'moloɣo] weak bond
	2	[i'pertata apo'fθeymata] supreme sayings		2	[ka'totato o'moloɣo] lowest bond
	3	[a'ðinata a'moɣa] weak anvils		3	[a'notato o'pliti] topmost private
	4	[ka'totata apo'fθeymata] lowest sayings		4	[p'rosfato opo'rona] new orchard
	5	[p'rosfata a'molutra] recent sand-bath		5	[θe'orato o'mometro] huge omometer
	6	[a'notata amorti'ser] best shock-absorbers		6	[θe'orato o'palio] huge opal

Table 5.1: The six items constructed for each of the vowel pairs under investigation.

Given that an extra condition (BC0) was added for the investigation of stop voicing for the previous chapter, BC0 items were also constructed for this part of the experiment. However, since the extra materials were added after the recording sessions had started, only the two male speakers (MS and CP) were still available and recorded producing these materials. For this reason, measurements from BC0 are not directly comparable with the other domains, but are only reported either in comparison within the two speakers, or for reference purposes. Table 5.2 shows the 6 items constructed for each vowel pair for this condition; they are all made up of a Definite Article followed by a Noun³.

³These constructions are only indirectly comparable to the results for BC0 from post-nasal stop voicing, since for that phenomenon a variety of negative particles, definite articles and temporal and conditional subordinate connectives were used as the pre-boundary words, and a Verb followed

/ao/	1	[ta o'moloya] bonds	/oa/	1	[to a'pokrifo] secret
	2	[ta opo'ra] bonds		2	[to a'poktima] asset
	3	[ta o'moixa] homophonous		3	[to a'poyevma] evening
	4	[ta o'mometra] omometers		4	[to a'povaro] dead-weight
	5	[ta o'palia] opals		5	[to a'moni] anvil
	6	[ta opti'ka] specs		6	[to a'molutro] sand-bath
/aa/	1	[ta a'povara] dead-weights	/oo/	1	[to o'moloyo] the bond
	2	[ta a'ponera] wake		2	[to o'palio] opal
	3	[ta a'moŋa] anvils		3	[to o'morizo] from same etymological root
	4	[ta a'pokrifa] secrets		4	[to o'moixo] homophonous
	5	[ta a'molutra] sand-bath		5	[to o'mometro] omometer
	6	[ta a'povlita] waste		6	[to o'monimo] the homonym

Table 5.2: The six items constructed for each of the vowel pairs for BC0.

5.3.1.2 Segmentation

The pre- and post-boundary vowels were segmented as one entity. The beginning of V1 and the end of V2 were marked in a separate Tier in PRAAT; this means that the V1V2 pair was segmented as one stretch, since it is not possible to identify where the boundary falls between the two vowels. Also, the duration of the pre- and post-boundary consonants was segmented, and again these values were extracted automatically by means of a script. The same segmentation technique as

the boundary. In this experiment a Noun followed the boundary, so that the same nouns were used as for conditions BC1-BC5 (only the neuter ones were used, since the masculine ones would not create a hiatus - the accusative of singular definite articles are [ton] ending in an /n/, thus stopping hiatus, while the neuter singular accusative of the definite article is [to] and the plural is [ta]). Furthermore, prosodically we do not expect a difference between a noun and a verb following the article, if they are placed in the same syntactic and phonological positions.

in preboundary durations was used for segmenting the closure and release of voiceless stops before and after the V1V2 pair. In the case of the first post-boundary consonant being an /m/, the spectrogram combined with the waveform were used for precise segmentations of the transition into and out of the nasal. Figure 5.1 shows an example of the segmentation of a wav file.

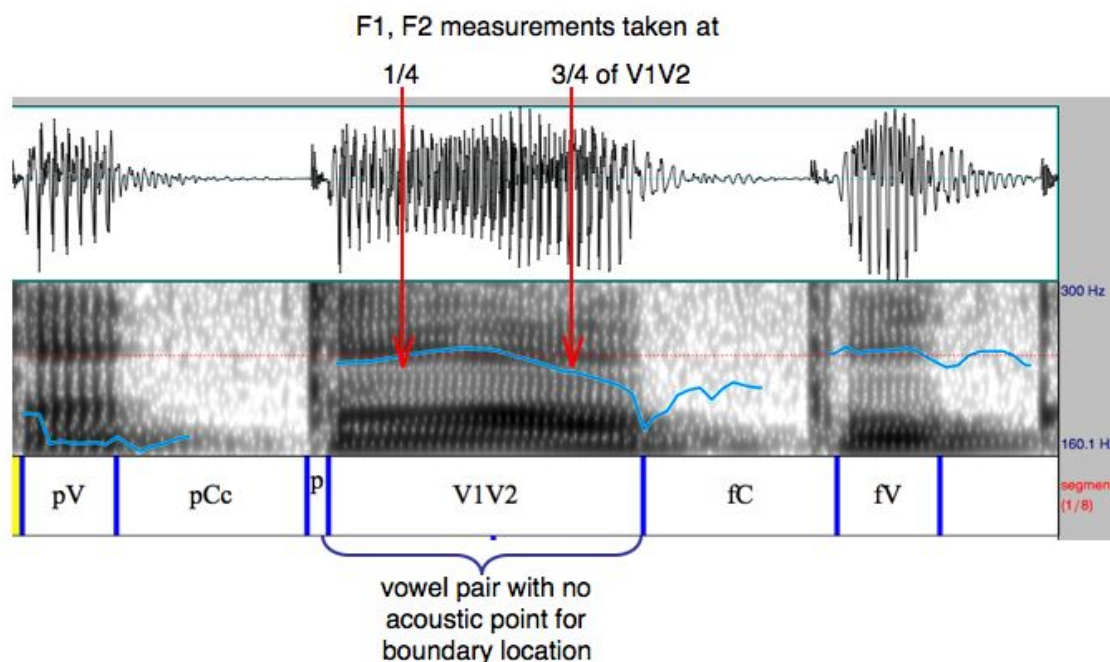


Figure 5.1: Example of segmentation of items constructed for the investigation of resolution of vowel hiatus. Sample shows the hiatus in V1V2, and the segmentation of the pre- and post-hiatus segments.

All data containing a silent interval between V1 and V2 were discarded, since they were not considered as representative samples of vowel hiatus resolution (123 sentences out of 240 total for BC5, suggesting that a large number of the sentences were excluded from this domain, and meaning that a pause was very common in BC5). Utterances were also discarded if there was a hesitation or disfluencies during the utterance (only two sentences were discarded for that reason, since during the recordings speakers were made to repeat sentences that had not been produced properly, and only very few were missed by the experimenter). Glottalisation has been found to mark prosodic structure in a number of languages, with more instances of glottalisation in higher order domains (Dilley et al. 1996). Glottalization in the present dataset was not common; there were only a few instances of creaky

voice separating the two vowels mainly occurring in BC5. The creakiness in voice in those instances appeared to act as a silence filler, and these items were also excluded from the dataset (five sentences in total). A total of 130 items were discarded out of 1248 items.

5.3.1.3 Measurements

In order to identify how vowel hiatus was resolved in our data, the following measurements were taken:

- the duration of the vowel pair (V1V2 duration)
- F1 and F2 values 1/4 and 3/4 into the V1V2 stretch (henceforth F1beg, F2beg, F1end and F2end for beginning and end measurements), allowing comparisons across vowel pairs and the identification of formant trajectories. Taking values from the beginning and end of the vowel pair provides a picture of the behaviour of formant trajectories. The measurement points were placed at 1/4 and 3/4 into the vowel pair to avoid influences from formant transitions into and out of the pair.

All measurements were performed using PRAAT. An adjusted version of Bert Remijsen's script⁴ was used for vowel formant calculation and extraction. This script computes formant values using Praat's 'To formant (burg)' algorithm, which allows the user to set parameters of gender etc, and it pauses allowing the user to check whether the measurements taken are correct and fix them manually (occasionally the algorithm would miss F1).

The F1 and F2 values were transformed to the ERB scale using Glasberg & Moore's (1990) formula:

$$ERB = 21.4 * \log_{10}[0.00437 * f + 1] \quad (5.8)$$

where f is the formant value in Hz⁵.

⁴<http://www.ling.ed.ac.uk/~bert/praatscripts.html>

⁵ERB is used in an effort to normalize as much as possible values across speakers, and is chosen based on Adank's (2003) investigation of which transformation algorithm is more reliable on the basis of a variety of analyses and tests.

5.3.2 Perception experiment

5.3.2.1 Procedure

The same three raters, as for the stop voicing perception experiment, were used. They were asked to classify all the items in the dataset; the baselines [oo] and [aa], and the non-identical pairs [ao] and [oa]. As in the case for stop voicing, the baseline pairs were used to make sure the accuracy was high for the cases that are unambiguous. The pre- and first two post-boundary syllables from the produced items were used as a context played back to the listener - not the whole items, again to avoid lexical bias. For example, from the item *αδύνατο απόγONO* [a'ðinato a'pɔɣONO], the syllables in bold letters were extracted. Figure 5.2 shows the part of the items that was extracted.

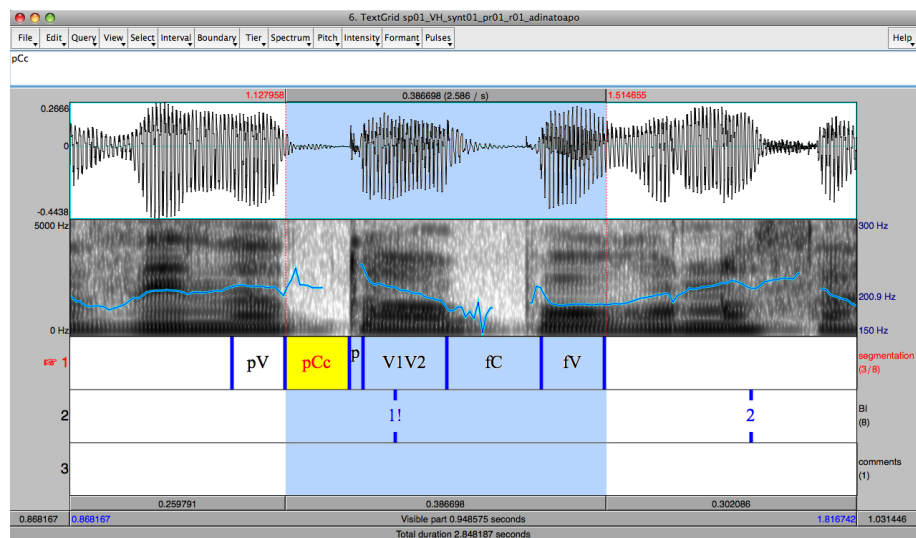


Figure 5.2: Highlighted in this screenshot is the section extracted from the item [a'ðinato a'pɔɣONO] (romanisation: *adynato apogono*). The vowel pair is [oa] and the preceding consonant and following syllable were extracted and played back to the raters (i.e. they would listen to “toapo”).

There were two parts to the experiment; in the first round the raters judged vowel pairs containing [ao] (and [aa], [oo] as baselines), and in the second vowel pairs containing [oa]. Half of the [aa] and [oo] instances were used for the [ao] experiment, and the other half for the [oa], ensuring no wavfile was heard twice. The experiment was again run in PRAAT (Boersma 2001), using an adjusted version of the MFC Identification script. Figure 5.3 shows the interface the raters used.

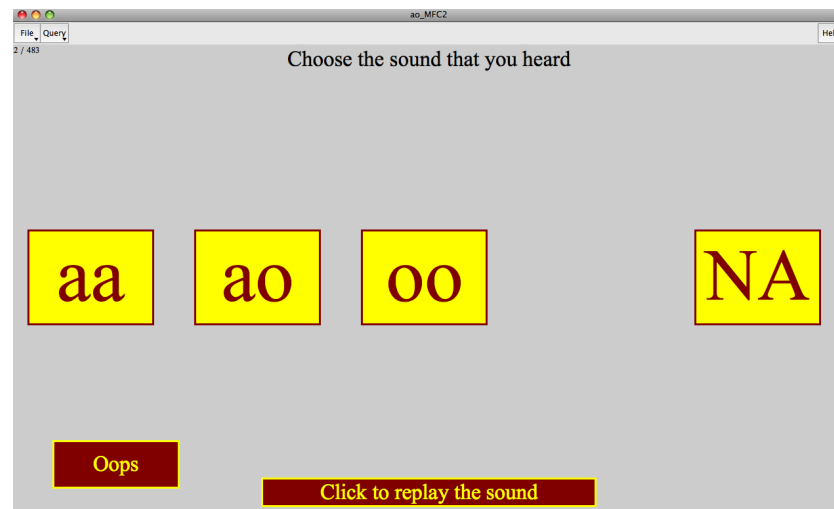


Figure 5.3: Screenshot of the interface for the perception experiment of the [ao] pair.

A Replay button was provided, and the rater was allowed to listen to the sound up to three times if needed to arrive to a conclusion. Again an ‘Oops’ button was provided, in case of a mistake, and a NA (No Answer) in case the rater could not reach a decision for any reason. None of the raters chose to use this option.

5.3.2.2 A note on the procedure and terminology

Using a perception experiment to classify the instances of vowel hiatus resolution in this experiment is regarded only as a tool to understanding whether impressionistic analyses of the phenomenon are able to capture its nature. For example, in post-nasal stop voicing it was seen that ‘nasal+voiceless stop’ sequences having undergone voicing were easily identifiable by listeners, supporting its apparent phonological nature. Moreover, the perception experiment provides initial insights into possible boundary strength effects by looking at how many times each item is classified as what. In the present experiment there are three available classifications for each vowel pair; [ao] can be classified as [ao], [aa] or [oo], and [oa] as [oa], [aa] or [oo]. In order to analyse the perception experiment we have termed each of those categories using phonological terms, that is the possible outcomes are grouped into three categories:

- **No assimilation:** instances of [ao] and [oa] classified as themselves - the output entails two vowels, which have formant values resembling those of the baselines at each endpoint. For example, in [ao] vowel hiatus, both V1 and V2 are present, and the beginning of the [ao] stretch resembles that of [aa], while the end that of [oo]
- **Deletion/Assimilation:** instances of [ao] and [oa] classified as either [aa] or [oo] - the output entails one vowel and the duration of the vowels can be used to identify whether deletion or assimilation has taken place with long pairs having undergone assimilation and short pairs deletion
- **Partial assimilation:** crudely represented by instances classified as no agreement - either of the two vowels' formants are in atypical positions. This category includes a phenomenon often referred to as “merger”, according to which the output of the hiatus is short in duration, but with atypical formant values (contrary to deletion/assimilation patterns).

It is important to point out that all of the above terms are used as working terms for classifying raters' answers. While they are grouped using phonological terms, it is not assumed that the phonetic differences across categories will necessarily be categorical. This classification is performed only in order to discuss how many times each resolution strategy appears in each boundary condition, and therefore compare whether perceptually there seems to be a categorical or variable boundary strength effect.

5.4 Results

5.4.1 Preliminary perception experiment

5.4.1.1 Inter-transcriber agreement

As for the previous chapter, the *kappa* statistic was first calculated testing the agreement between transcribers. The agreement reached lower levels than for the identification of stop voicing in the previous chapter: the kappa value was at $\kappa = 0.73$ for [ao] and $\kappa = 0.68$ for [oa]. The raters were called to classify the baseline vowel pairs [aa] and [oo] as well as the non-identical vowel pairs [ao] and [oa]. This ensured that the unambiguous cases had been classified correctly with

little inter-transcriber disagreement. Results showed that indeed in all boundary conditions almost all instances of [aa] were classified as [aa] and [oo] as [oo] with very few instances of no agreement. This ensured that all three raters were able to perceive the baseline conditions, and agreed in their classifications. This suggests that, if the formant trajectories are clear, as they are in the case of the baselines, the task of identifying the vowel pairs is very easy. Moreover, the low inter-rater agreement for non-identical vowel pairs seems to mirror the fact that resolution of non-identical vowel pairs might not fall into neat phonetically distinct categories which are easily perceived. Similarly, this low kappa value is different to the high inter-transcriber agreement found in the previous chapter (see previous chapter for discussion of acceptable levels of kappa values, and for comparison to values found in stop voicing identification). This is a first indication that the phonetic output of the vowel hiatus resolution strategies ranges in a continuum of not easily identifiable values when placed out of their lexical context.

These results initially suggest that the task is a rather difficult one, and that using perceptual classifications for identifying the phonetic output of vowel hiatus, as has been done extensively in traditional research, might not be the most appropriate way of addressing the phenomenon. Moreover, it shows that there must be many instances of partial assimilation; if the phonetic characteristics of either one of the two (or both) vowels are atypical, it is reasonable to assume that raters will have difficulties classifying those instances with high agreement, especially in a perception experiment where lexical information is stripped from the input.

Table 5.3 shows the percentage times each vowel pair was classified as either one of the available categories. It is evident that ‘no assimilation’ and ‘no agreement’ are the most common cases with 47% of the [ao] instances and 40% of the [oa] instances classified as themselves, and 32% of the [ao] and 49% of the [oa] instances not having reached agreement. Given that it is a forced choice experiment, it is anticipated that the ‘no assimilation’ instances would be quite high, since the raters were at least trying to provide an answer. While the raters did have the option of a ‘No Answer’ button, they did not use it, but rather tried to provide an answer in almost all occasions. It seems that, given the large number of disagreement, they were comfortable with providing an answer classifying sounds about which

maybe they weren't completely certain. This is an added example of how impressionistic analyses might force the classification of gradient phonetic outputs into predefined categories. It is interesting to note that there are very few instances of V1 and V2 deletions/assimilations, contrary to previous analyses of the phenomenon which assumed that the most common resolution strategy is V1 deletion. Deletions/assimilations of [oa] seem somewhat more common than for [ao], while also V2 deletion/assimilation again seems more common (more [aa] classifications than [oo] for [ao], and conversely more [oo] than [aa] for [oa]). The large number of 'No Agreement' suggests that there must be instances of intermediate phonetic realisations, which are hard for the raters to classify.

Thus, the data support the impression that partial assimilation is very common in the dataset, and add that many instances could potentially be identified as having undergone no assimilation, if forced to reach a decision.

	ao	aa	oo	No Agr.
ao	47.3	10.7	9.8	32.2
	oa	aa	oo	No Agr.
oa	40.5	2.6	13.6	42.9

Table 5.3: Mean percent of how many times [ao] and [oa] vowel pairs were classified by the raters as themselves, or as either of the baselines pooled across prosodic conditions.

5.4.1.2 Boundary strength effect

Table 5.4 and Figure 5.4 show the percent of times each vowel pair was recognised as either one of the vowel pairs, as well as the percent of the three raters not having agreed (either one rater disagrees with the other two, or all three have given different answers).

First, it can be seen that there is a substantial amount of disagreement among raters (as was seen from the kappa statistic), supporting a view of partial assimilation being the most common resolution technique, since it is one that cannot be identified perceptually as such. Second, the trend for a gradient effect of boundary strength can be seen in all classifications; the amount of disagreement between raters is gradiently lower in higher domains, and the amount of no assimilation is higher

	ao				oa			
	ao	aa	oo	NA	oa	aa	oo	NA
BC1	27.3	18.2	18.2	36.4	30.2	2.3	2.3	62.7
BC2	25.6	11.6	9.3	53.5	11.1	8.9	31.1	48.9
BC3	32.6	23.3	11.6	32.6	24.4	4.4	31.1	40
BC4	68.9	11.1	6.7	13.3	52.1	-	14.6	33.3
BC5	79.5	-	-	20.5	75.6	-	2.2	22.2

Table 5.4: Table showing the percentage values of how many times the vowel pairs [ao] and [oa] are classified as themselves or as one of the baselines for each prosodic condition across speakers ($n \approx 60$ for each boundary condition for each vowel pair). The NA columns stands for the percentage rate that the raters did not reach an agreement (No Agreement).

in higher domains. This shows that more “canonical” (not deleted/assimilated or partially assimilated) instances exist in higher domains. Second, there are gradually fewer instances of deletions across boundary conditions, with almost no instances in BC5 (however do keep in mind that very few instances were used in BC5). It can also be seen that instances of ‘no agreement’ and ‘deletion/assimilation’ are not categorically blocked in any domain, thus excluding an analysis of a categorical effect from boundary strength.

Finally, there seem to be more occurrences of deletion/complete assimilation for [oa] and most of them are classified as [oo] signifying a V2 deletion tendency for the vowel pair. For [ao] again there are marginally a few more instances of [ao] classified as [aa], again showing a preference for V2 deletion/assimilation. However, both vowel positions seem to be able to undergo deletion/assimilation in Modern Greek.

A separate analysis was performed for the two speakers that had also produced data from BC0, and the pattern was similar to BC1 with very small percentage changes towards more instances of classified [ao] and [oa] as [aa] and [oo], supporting the trend found.

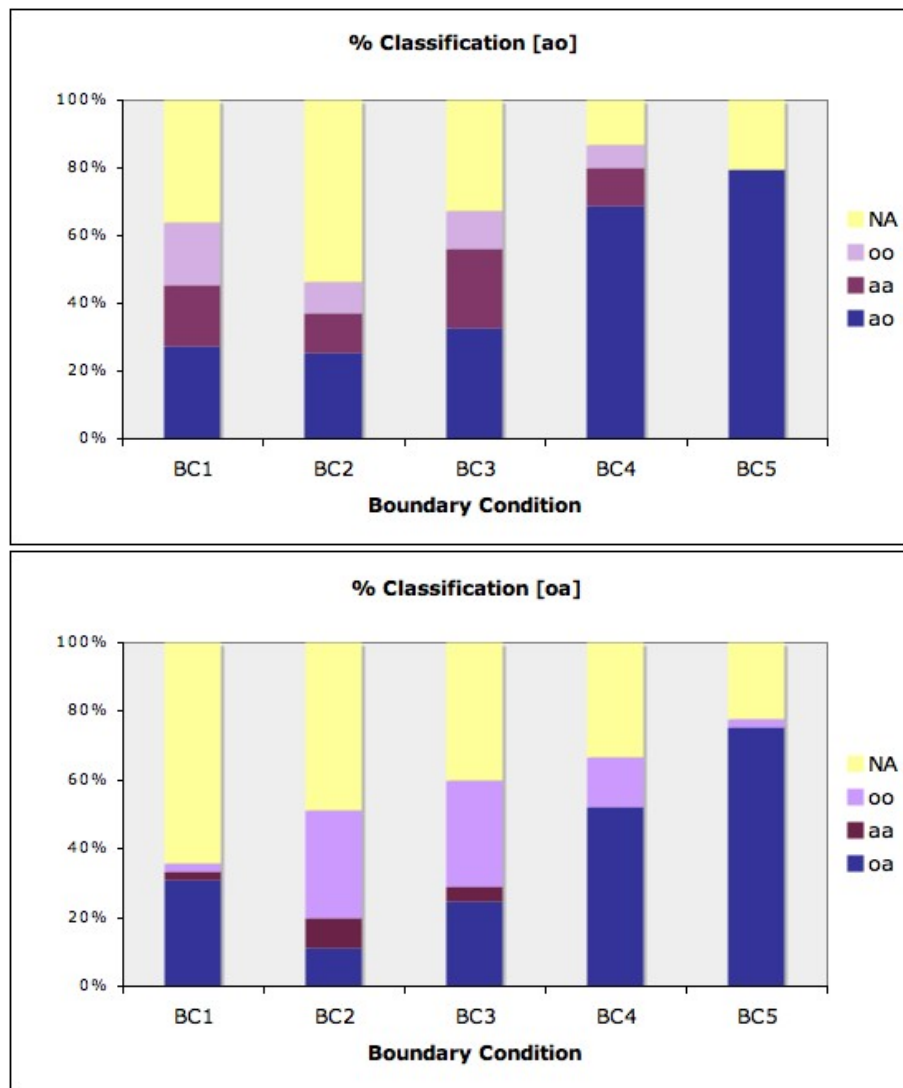


Figure 5.4: Percent classification of [ao] and [oa] for each prosodic condition by pooling all raters' results. BC0 is not represented since only two speakers produced instances for that condition.

5.4.1.3 Conclusions on perception experiment

The difficulty raters were faced with in classifying vowel pairs was apparent in their lack of agreement, suggesting that traditional impressionistic analyses using perceptual classification of vowel pairs into phonologically predefined categories needed instrumental support for their claims. Additionally, the low inter-rater agreement suggests that there must be many instances of variable phonetic output, which are therefore difficult for the raters to classify. Boundary strength effect

seemed to influence the distribution of classified resolution strategies variably, since no categorical effects were apparent. Moreover, there seemed to be gradually fewer instances of deletion/assimilation (that is non-identical vowels classified as [aa] or [oo]), and more instances of ‘no resolution’ (that is [ao] and [oa] pairs classified as themselves) higher in the structure. This suggests that there is less coarticulation higher in the structure, but the existence of deleted and partially assimilated items in all conditions shows that there is no categorical effect.

The following section will investigate the actual phonetic implementation of vowel hiatus resolution. First we test whether the phenomenon seems to be coarticulatory, and then we test the effect of boundary strength on it.

5.4.2 Production experiment

5.4.2.1 Sandhi or coarticulation?

The hypothesis that the phonetic output of vowel hiatus resolution is gradient is tested. The beginning and end points of vowel pairs [ao] and [oa] were compared to those of baselines [aa] and [oo]. The expectation is that, if no resolution is taking place and both vowels are present, measurements of F1 and F2 formants for [ao] will resemble those of [aa] at the beginning and those of the [oo] at the end of the formant trajectories, and the other way round for vowel pairs [oa]. If V1 deletion/complete assimilation takes place as would be expected by impressionistic traditional phonological analyses, F1 and F2 formants for V1 of V1V2 pairs will resemble those of V2V2 formants at the beginning and end of V1V2, and if V2 deletion takes place, formants of V1V2 at the end of the trajectories will resemble those of V1. If, finally, partial assimilation is taking place, the formant values of V1V2 will fall between those for V1V1 and V2V2.

Four One-way ANOVAs were performed, one for each dependent variable (F1beg, F1end, F2beg, F2end), with the vowel pair as the dependent variable (four levels). The effect of vowel pair was significant on all dependent variables (F1beginning: $F(3, 943) = 110.046$, $p < .001$, F1end: $F(3, 940) = 75.179$, $p < .001$, F2beginning: $F(3, 942) = 179.728$, $p < .001$, F2end: $F(3, 745) = 171.605$, $p < .001$). Table 5.5 shows the mean ERB values of the four vowel pairs for all dependent measurements,

pooling their values across prosodic conditions. The last row shows Bonferroni post-hoc tests testing the differences across vowel pairs within each dependent variable.

	ERBF1beg	ERBF1end	ERBF2beg	ERBF2end
ao	13.04	11.92	17.83	17.02
oa	11.41	12.42	16.71	17.54
aa	13.65	13.63	18.14	18.31
oo	11.03	10.79	16.44	16.37
Post-Hoc	aa>ao>oa=oo	aa>oa=ao>oo	aa>ao>oa>oo	aa>oa>ao>oo

Table 5.5: Mean ERB values for all vowel pairs in all dependent variables, pooled across boundary conditions. Comparisons across vowel pairs are shown with Bonferroni post-hoc tests in the last row.

Figure 5.5 shows a graph of all the measurement points for each vowel pair. The Bonferroni tests and Figure 5.5 show the following regarding the general trends of hiatus resolution in the data:

- vowel pairs [aa] and [oo] assume the highest and lowest points respectively in F1 and F2, and remain in the same range from the beginning to the end of the trajectories, and therefore provide clear baselines
- vowel pair [ao] has higher values than [oa] at the beginning of the trajectories, and lower at the end, as anticipated in typical formant trajectories
- values for both [ao] and [oa] fall in between those of [aa] and [oo] showing that most cases must be undergoing partial assimilation (with the exception of [ao] which is similar to [oo] at the end of the trajectories). However, it is important to note that these distributions could also arise if they only entailed the phonetic output from two categorical strategies, for example, complete assimilation and no assimilation. To show that this is not the case, and that partial assimilation is indeed the most common resolution strategy, Figures 5.6, taken from speaker OM, show F2 trajectories for all instances of that speaker producing vowel pairs [ao] and [o]. It is shown that indeed partial assimilation is very common (all speakers had similar graphs)
- the upper and lower ends of the boxplots for both [ao] and [oa] overlap substantially with their baseline counterparts. For example, F1end of [ao] has

instances that overlap with those of [aa] instead of the typically anticipated [oo], suggesting that, although partial assimilation seems to be the most common resolution output, there must exist instances of no deletion/assimilation and of complete assimilation/deletion in the data.

In conclusion, this section provided support for recent research proposing that non-identical vowel hiatus resolution in Modern Greek is a coarticulatory phenomenon.

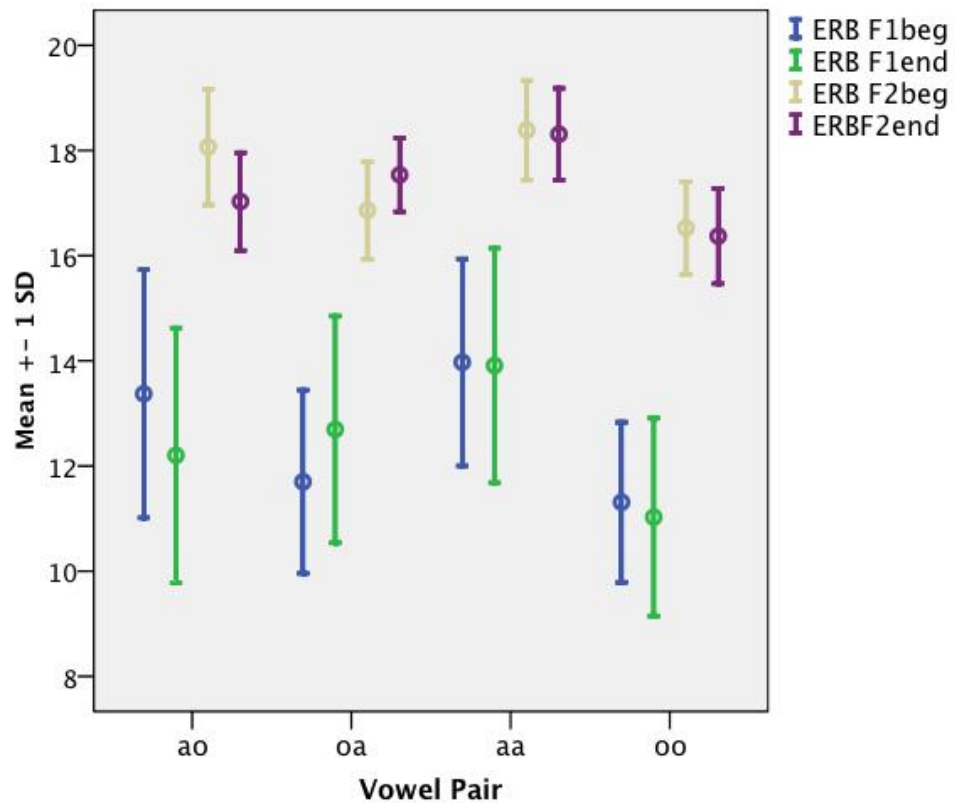


Figure 5.5: Mean F1 and F2 values (ERB) at the beginning and end of formant trajectories (and $\pm 1SD$) for each vowel pair.

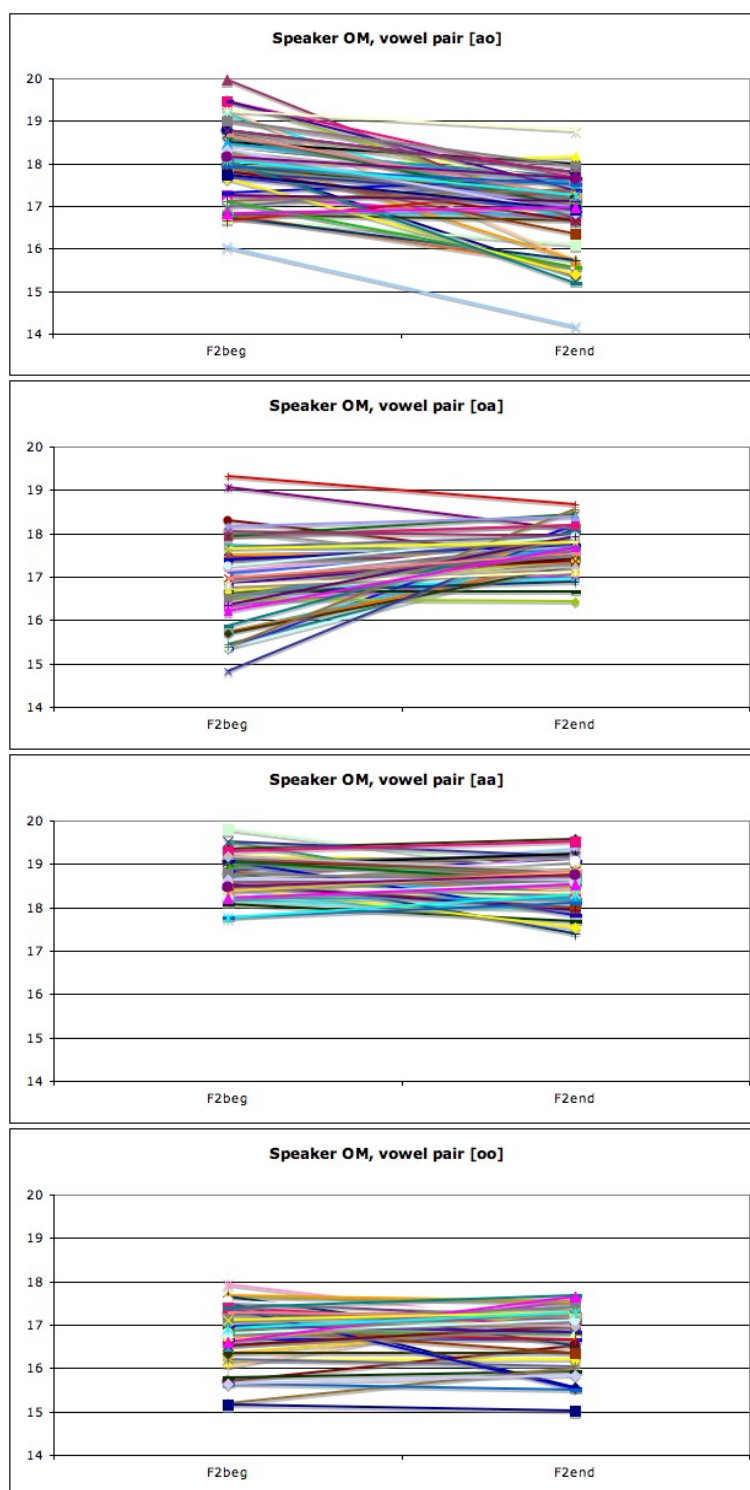


Figure 5.6: F2 trajectories for all instances of [ao], [oa], [aa] and [oo] vowel pairs for speaker OM. This graph exemplifies the gradient phonetic output of vowel hiatus resolution in Modern Greek.

5.4.2.2 Qualitative or variable boundary strength effect?

The effect of boundary strength is tested using four ANCOVAs on F1beg, F1end, F2beg and F2end ERB values for [ao] and [oa], with the duration of V1V2 stretch as the covariate. The independent variable was the boundary condition with five levels (BC1-BC5). Duration was not found to be a significant covariate for any of the dependent variables. Two possible effects are expected to occur; either the effect of boundary strength on the application of hiatus resolution exercises a qualitative effect allowing all sorts of resolution outputs within a specific domain and blocking them in all other domains only allowing each vowel of the vowel pair to assume its ‘typical’ position, or a variable effect, according to which all sorts of resolution strategies are allowed across all types of boundaries, but the values of the formants getting gradually more extreme (closer to their ‘typical’ values) suggesting fewer instances of gradual phonetic outputs, and more instances of typical outputs at stronger boundaries.

Figure 5.7 shows an approximation of how formant trajectories for F2 [ao] vowel pairs are expected to be realised if there is a qualitative or a variable boundary strength effect. The first graph shows how formant trajectories are expected to be if there is no resolution and both vowels assume their ‘typical’ values. The second graph shows the expected finding if there is a categorical effect separating lower from higher domains. The line named ‘qualitatively distinct domain’ depicts the lower domains where coarticulation is allowed, and therefore the formants are not in their typical positions. The line named ‘all higher domains’ shows the domains where coarticulation is blocked, and therefore formants assume their typical positions. Finally, the third graph shows the type of formant trajectories expected if the effect is variable. It is seen that there will be instances ranging from no coarticulation to partial assimilation. Deletion and full assimilation have not been depicted in order not to clutter the diagram, but they are also possible and would be plotted as trajectories starting from an [a] and ending in an [a] position for V2 deletion, and similarly starting from an [o] and ending in an [o] position for V1 deletion. F2 [ao] formant trajectories are used as an example.

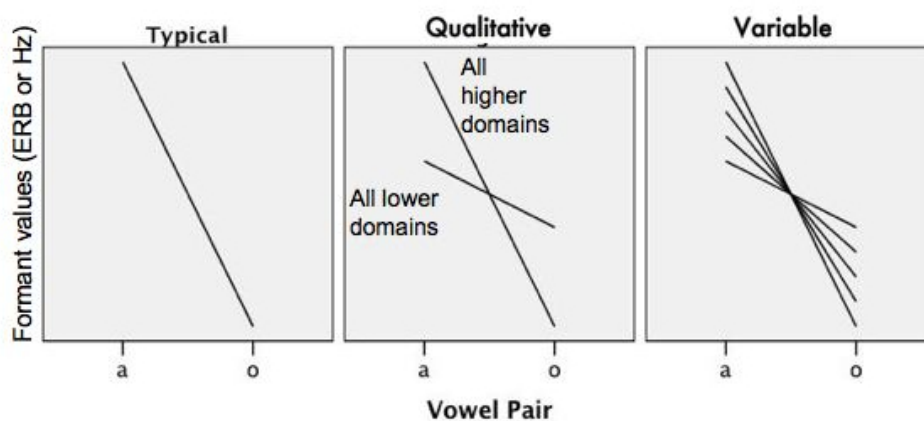


Figure 5.7: Presentation of expectations from boundary strength effect; the first panel shows how formant trajectories of F2 [ao] vowel pairs are expected to look when they meet their typical positions (baseline [a] or [o] positions). The second figure shows a categorical effect whereby one condition is statistically different to the others, within which there is coarticulation allowed, and the third figure shows a gradient effect across domains, whereby more extreme formant values are found higher in the hierarchy, but the effect is not categorical.

Boundary strength did not significantly affect the formant values of either [ao] or [oa] either at the beginning or the end of the trajectories. This means that the way vowel hiatus was resolved was not influenced significantly by boundary strength. Table 5.6 shows the mean ERB values for all formant measurements for [ao] and [oa], and Figure 5.8 shows the formant trajectories from the beginning to the end of V1V2 for vowel pairs [ao] and [oa] for each prosodic condition. This figure depicts the part of the vowel space zoomed in to the positions of [a] and [o]. Values for [a] and [o] marked with blue letters are those provided by Fourakis, Botinis & Katsaiti (1999), and are included solely for reference (since our design only contained a and o vowels, a vowel space could not be created on the basis of solely those values); similarly the black dot at the far left mid point of the graph shows the position of an imaginary schwa, as the mid point of the vowel space. One interesting issue to point out is that, although not significantly different, BC4 has marginally more extreme values for both vowel pairs, initially suggesting that fewer instances of partial assimilation may be included in that condition. BC0 and BC5 are not represented in this graph for different reasons; BC0 only contains results from two

speakers, thus the trajectories will not be comparable to the other conditions, and BC5 entailed many instances of silent intervals between V1 and V2, leaving only a few instances for analysis. Separate analyses for the two speakers including BC0 again showed no effect.

This result shows that there is a lot of variability within each prosodic condition, as was seen in Figure 5.6 in the previous section, where all F2 instances from speaker OA for all vowel pairs were plotted. That figure provided a clear insight into the fact that there is variability in the data. That figure did not separate across boundary conditions, therefore it could have been the case that all instances with no resolution could have come from one domain, and all instances with most coarticulation could have come from another. However, this is not the case, and boundary strength does not seem to affect formant positions.

		BC1	BC2	BC3	BC4	BC5
[ao]	F1beg	13.2	13.2	13	13.3	12.1
	F1end	12.3	11.7	12	11.8	11.6
	F2beg	17.7	17.9	17.8	18.1	17.4
	F2end	16.1	16.9	17.2	16.9	16.9
[oa]	F1beg	11.5	11.4	11.6	11.2	11
	F1end	12.5	12.5	12.4	12.5	11.8
	F2beg	16.8	16.9	16.8	16.4	16.3
	F2end	17.6	17.6	17.5	17.6	17.1

Table 5.6: Mean ERB values from all [ao] and [oa] vowel pairs in all prosodic conditions. The effect of boundary strength was not significant for all any measurement.

The effect of boundary strength was also measured on the duration of the vowel pairs. It was tested using two one-way ANOVAs, one for [ao] and one for [oa] with boundary condition as the independent variable with five levels (BC1 to BC5). It was found that the effect was significant for both vowel pairs ([ao]: $F(4, 245) = 8.065$, $p < .001$, [oa]: $F(4, 236) = 6.599$, $p < .001$). Bonferroni post-hoc tests showed that vowel pairs in BC4 were longer than BC1, BC2 and BC3. Table 5.7 shows mean values for all vowel pairs in each boundary condition⁶. Figure 5.9 shows that there is substantial variation in the duration of the vowel pairs within boundary conditions.

⁶BC0 is again not reported in these results, since only two speakers were recorded producing these instances. Separate analyses of these two speakers revealed no differences between BC0 and BC1, although the mean durations for BC0 were slightly longer than BC1, BC2 and BC3.

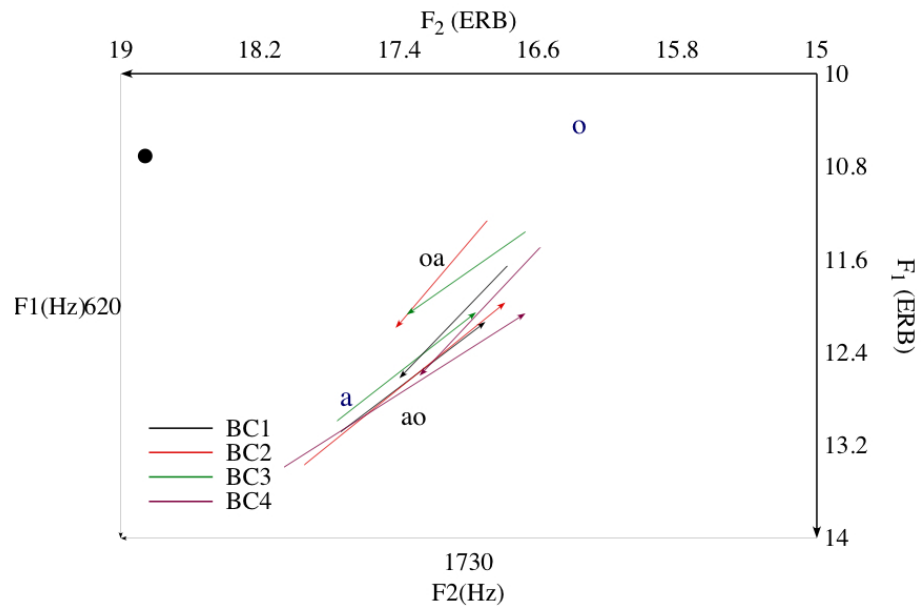


Figure 5.8: Trajectories from beginning to end of the [ao] and [oa] vowel pairs for conditions BC1 to BC4.

There is a clear effect of the length of the utterance (BC4) on the duration of the vowel pairs, suggesting that longer utterances give rise to longer durations, as was found in Chapter 3, and the effect is apparent in both vowel pairs⁷.

	Boundary Condition					Post-Hoc
	BC1	BC2	BC3	BC4	BC5	
ao	105	104	103	134	124	123<4
oa	99	97	100	124	116	123<4
aa	86	102	117	136	125	1<345,4<2
oo	96	99	101	114	121	1<5

Table 5.7: Mean duration (sec) of all vowel pairs in all prosodic conditions. The effect of boundary strength was significant for all vowel pairs, and significant Post-Hoc Bonferroni results are reported in the last column.

One thing to note is that BC5 is shorter than BC4 for both vowel pairs. Although this difference is not significant, it is consistent, and it is due to the small number

⁷The baseline vowel pairs are not used in this section because identical vowel hiatus resolution is different to non-identical vowel hiatus resolution. For example, identical hiatus can be resolved by degemination (Nespor 1987), and has been found to result in more deletions than non-identical hiatus (Fallon 1994). Thus, although formant values are not influenced by the different strategies used for identical-vowel resolution (since the formant trajectories remain in the same position), the durational patterns might differ.

of instances remaining for analysis in BC5 after discarding instances with pauses and creaky voice. This is taken to be due to the small number of BC5 items comparatively to BC4, added to the fact that, in order for BC5 not to have a pause (recall the construction involves two sentences separated by a fullstop), it means that the particular rendition of the sentence must have been rather fast, therefore probably making the production of the segments close to the boundary shorter than in BC4.

Finally, it seems that there are instances of fully assimilated and deleted vowels in the data. Formant values are not expected to differ across processes (i.e. assimilation and deletion); e.g. V1 assimilation in V1V2 results in V2V2, and V1 deletion results in V2, meaning that both results will have the V2 formant characteristics. However, assimilation vs deletion can be identified on the basis of durational patterns. Deleted items will be short, resembling singletons, and assimilated patterns will be long, resembling the duration of two vowels. In Figure 5.9 it can be seen that there is large variability in the duration, ranging from very short values indicating deletion, to very long values indicating either full assimilation or production of both vowels⁸

The present section, by investigating the acoustic manifestation of vowel hiatus pairs in a variety of boundary conditions, has clearly shown that boundary strength does not influence vowel hiatus resolution in a categorical manner, but rather it was shown to affect boundary strength in a gradient manner, with longer durations in higher domains, and less instances of deletion/assimilation/no agreement classifications in higher domains. The effect was not at all pronounced in the actual formant measurements, showing that it is not a very clear effect, and definitely not one that can be used for supporting the existence of labelled domains. The difficulty raters had in agreeing on the classification of vowel pairs indicated that traditional impressionistic analyses using perceptual classifications need revisiting, since the

⁸The issue of whether there is assimilation and deletion in the data was also tested using results from the perception experiment. Instances of [ao] and [oa] classified as [aa] or [oo] are believed to have undergone either deletion or assimilation. The duration of those instances was plotted and found to range in a continuum of values from short to long indicating (i) that both types of resolutions take place in the data, and (ii) that they are not phonological since they did not give rise to a binary distribution separating the deleted from the assimilated instances. Also, duration did not covary with formant values (p. 157).

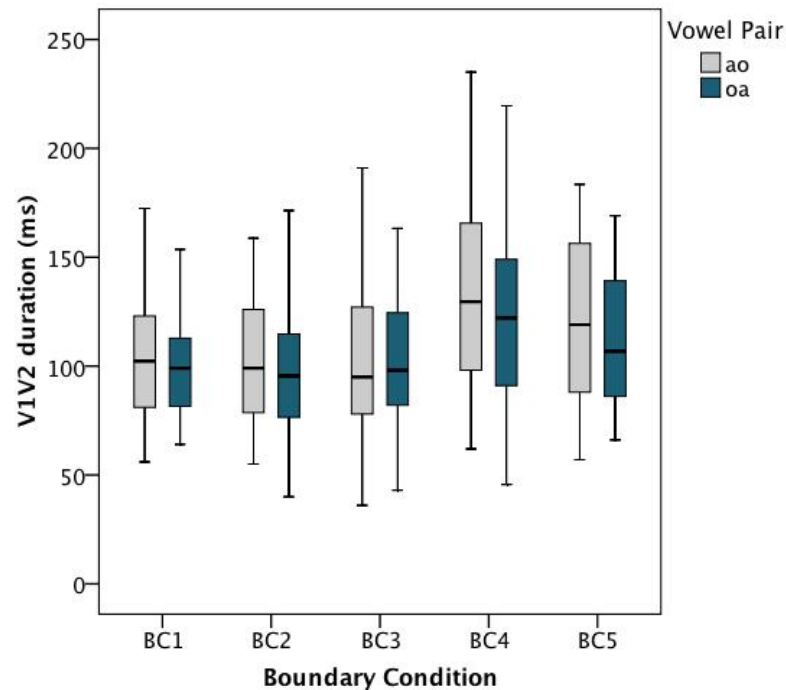


Figure 5.9: Mean duration (sec) for all vowel pairs in all boundary conditions.

phenomenon seems to be coarticulatory and cannot fit into predefined phonological categories. The amount of coarticulation across boundaries seemed to be influenced very marginally, and the effect was mainly apparent on durational patterns - not formant patterns.

5.5 Conclusions

Vowel hiatus resolution in Modern Greek has been claimed by different studies to entail either a phonological change resolved with V1 deletion (Kaisse 1985, Nespor & Vogel 1986), or to be a phonetic coarticulatory process resulting in a continuum of phonetic outputs (Pelekanou & Arvaniti 2001, Baltazani 2006, Kainada 2007, Arvaniti 2007). The present study has provided support for the recent claims that the phonetic output of vowel hiatus resolution is a phonetic coarticulatory process using a perception and production study. Furthermore, it was shown that in the present dataset partial assimilation was the most common output of formant

trajectories. Deletion and full assimilation were also apparent, but all these classifications were shown to range in a continuum of possible values, and not in distinct categories, supporting the view that the phenomenon is the result of coarticulation.

The main goal of the present experiment was to test the effect of boundary strength on the resolution of hiatus. The motivation for this question stemmed from conflicting accounts from both traditional impressionistic analyses of the phenomenon, as well as from more recent phonetic analyses supporting that the effect is qualitative, but disagreeing with respect to the prosodic domain blocking its application. It was shown that the effect of the boundary strength on the actual raw data is not qualitative; rather, it was shown that, if an effect was apparent, it was mainly apparent on the duration of the vowel pairs, while the formants across prosodic conditions were not influenced significantly. This was supported by looking at how listeners classified all items; the distribution of classifying [ao] and [oa] vowel pairs across conditions as themselves or as either of the baselines was variable, with more instances of no resolution, less instances of ‘deletion/assimilation’ (i.e. classification as a baseline vowel), and less instances of ‘no agreement’ higher in the structure. Thus, the effect of prosodic boundary strength on the resolution of this phenomenon is taken to be variable, contrary to previous analyses (Nespor & Vogel 1986, Condoravdi 1990, Pelekanou & Arvaniti 2001). Moreover, these results highlight how the mismatch found in recent studies by Pelekanou & Arvaniti (2001) and Baltazani (2006) according to which the effect was reported to be categorical but to distinguish different domains is resolved in this study, since it is shown that by using a variety of boundary strength conditions we can conclude that the effect is variable.

This chapter has mainly shown that vowel hiatus resolution is coarticulatory above the word level. Findings from BC0 from the two speakers suggested that there was no qualitative effects on that condition either. However, especially within this lower condition it can be proposed that some type of phonological rule of deletion might be taking place, like the one proposed by Nespor & Vogel (1986) for deletion between [ta 'exo] → [taxo], which native speakers of Greek agree that it seems like the vowel is completely deleted. However, the present chapter has not specifically tested this lower domain. It seems that a variety of vowel pairs will need to be

investigated using a variety of stress patterns and parts of speech, before concluding whether there is possibly some type of boundary strength effect there that has not been addressed in this chapter. However, in higher domains this chapter has clearly shown that boundary strength effect is variable, and cannot offer support for labelled domains, contrary to Condoravdi's analysis.

The present results relate to a number of issues in the Greek literature for vowel hiatus resolution, which we did not set out to investigate, but which are nevertheless interesting. First, it has been shown that V1 deletion is not the most common output when resolving vowel hiatus in Modern Greek, an assumption that has long been held within traditional analyses of the phenomenon. Rather, in line with Baltazani's (2006) and Pelekanou & Arvaniti's (2001) analyses, we found that deletion and assimilation are possible phonetic outputs as the extreme of a number of phonetic gradual productions, and that they are definitely not that common. Moreover, the position of the vowel within the vowel pair did not seem to be important in the present dataset, since no durational differences were found between instances classified as [ao] by the raters vs those that were classified as [oa]⁹.

It was found that the duration of the vowel pairs was not a significant covariate for formant values. Both Zsiga (1997) and Baltazani (2006) reported on the duration being a significant covariate, supporting the analysis of the phenomenon as the result of gradient articulatory overlap, according to which the more gestures overlap the more the formant values are blended, and the shorter the durations of V1V2 are. No articulatory data were gathered for the present experiment, thus no conclusions can be drawn regarding the articulatory representation of the phonetic output. However, the lack of correlation means that the results offer no support for the gestural overlap account, and requires further investigation.

Finally, it was found that the influence of boundary strength on the phenomenon more closely relates to the type of influence exerted by boundary strength on pre-boundary durations than to that of post-nasal stop voicing, which was found to be influenced qualitatively by boundary strength. This gives rise to a difference

⁹This analysis was not reported in the text, since it only links with different theoretical questions with Modern Greek vowel hiatus resolution.

between the application of the two purported sandhi processes. These findings provide support for the hypothesis that maybe phonetically gradient processes are more prone to be influenced variably by boundary strength, like pre-boundary durations, while sandhi processes are more prone to signal constituency in a qualitative manner. This hypothesis will be further discussed in the Discussion chapter, where also findings from the intonational marking of constituents will have been gathered, and implications of these findings for theories of prosodic structure will be discussed in light of all the phenomena investigated.

CHAPTER 6

F0 alignment and scaling

The question whether the prosodic structure entails labelled domains of distinct type, as proposed by the labelled view, is further investigated in this chapter by looking at the acoustic realisation of intonation contours, the last acoustic correlate of prosodic structure to be investigated in this thesis. The goal of this chapter is to test whether scaling and alignment patterns of intonation contours in the vicinity of varying boundary strength contexts are influenced in a manner supporting the existence of labelled domains. The experiment is meant to test whether and how intonation contours change across conditions to provide support for discrete differences across boundary conditions.

6.1 Introduction

6.1.1 Intonation contours, prosodic constituents and gradience in intonation

The present chapter will focus on research that has mainly used intonation contours for the postulation of prosodic domains (Beckman 1986, Pierrehumbert & Beckman 1988, Beckman 1996, Jun 1998), since different studies have focused on different acoustic correlates to prosodic structure.

Intonation contours at the vicinity of boundaries have been crucial in the identification of prosodic domains within the labelled view, as was seen in the Introduction to this thesis. One of the main proposals within the analysis of intonational events and prosodic structure has been that there is a direct link between intonational

events and prosodic domains. That is, intonational events identified at the edges of constituents were proposed to be in a one-to-one relationship with the number of prosodic domains in the structure (Pierrehumbert & Beckman 1988). This work built on previous research by Pierrehumbert (1980), who proposed a system for the analysis of the intonational phonology of English, whereby pitch contours consist of linearly distributed discrete intonational entities, which associate with the segmental string. The association is regulated by the prosodic structure, and refers to heads and edges, and as such relies crucially upon the prosodic structure and the definition of prosodic constituents. The link between intonation and structure was clearly proposed by Beckman & Pierrehumbert (1986) and Pierrehumbert & Beckman (1988), and was further elaborated with the emergence of the ToBI system (Silverman et al. 1992). In its initial form, the ToBI system only described how intonational entities would associate with respect to the prosodic structure, keeping the intonational tier and the prosodic structure tiers clearly separate.

This separation was clear at the beginning of the emergence of the Tone and Break Index annotation system; the Tone tier entailed the analysis of the intonational events of the utterance, and the Break Index tier marked the boundary strength of the disjuncture between each orthographic word. Seven levels were available to the annotator in the Break Index tier, reflecting the understanding that differences across levels are gradiently varying in terms of their relative strength (Price et al. 1991). However, with the emergence and widespread use of the ToBI system (which has been extended to many languages including Greek, Dutch, German, Korean, Japanese etc.), researchers adjusted the theory by indirectly subordinating the Break Index analysis (which mainly represents prosodic grouping) to the Tone analysis in proposing that the number of boundary strength levels represented in the Break Index is linked to the number of intonational entities marking the domains. By that it is meant that there are as many Break Index levels as many distinct intonational entities in the language, and, importantly, each of those levels are not only differing in terms of their relative strength, but mainly are different types of prosodic domains.

Therefore, throughout the years the proposal put forward has settled into a schema that can be summarised as the one seen in Figure 6.1, proposing the existence of

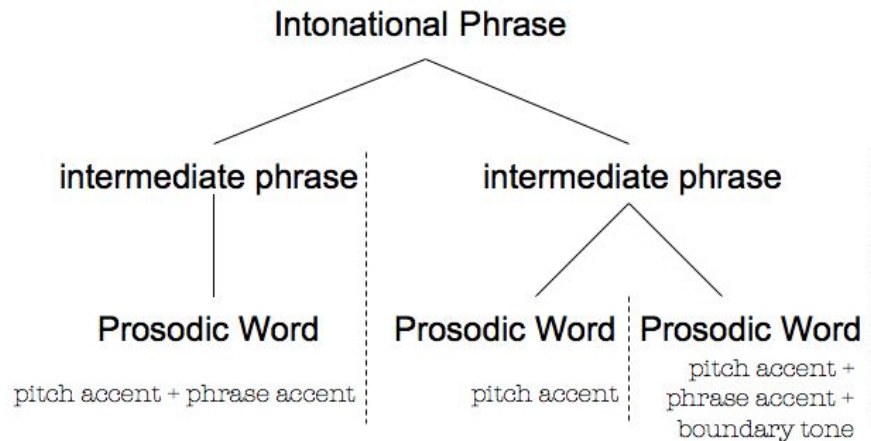


Figure 6.1: Tree depicting prosodic structure, as expected by current analyses of the Intonational Approach.

three types of intonational events, namely PITCH ACCENTS, PHRASE ACCENTS and BOUNDARY TONES. Pitch accents associate with Prosodic Words (or Accentual Phrases), and edge tones are broken down to (i) phrase accents signalling intermediate phrases, and (ii) boundary tones which signal intonational phrases. The edge tones are in theory different to pitch accents:

“Pitch accents associate with stressed syllables within their phrasal domain[. . .] Boundary tones associate with the boundary for which they are diacritically marked” (Hayes & Lahiri 1991, pp. 54).

There are two issues with the way this system has evolved that affect our understanding of prosodic structure, and we will turn to each of those now. Both have to do with the assumption inherent in the current view of the intonational approach that there are as many prosodic domains as discrete intonational entities.

The first issue relates to boundary strength, and in particular to the issue of whether there are possibly more boundary strength gradations available in the acoustic manifestation of intonation than are captured by the current analyses in labelled views of intonational approach. In the current state of the intonational approach there are

three distinct tonal events which are associated with three prosodic domains¹. For example, when a researcher marks a disjuncture in a ToBI analysis as a level 3, it is immediately assumed that the associated tone will be a phrase accent, and that the relevant prosodic domain is one of intermediate phrase. Subordinating the number of available break index gradations to the number of identified intonational events has not been the result of phonetic analysis, but served as an expedient means for ToBI analyses. In particular, this proposal presupposes the existence of a set of labelled domains within the prosodic structure, and discards the information that was inherent in the initial Break Index analyses, which stated that there can be up to seven gradations (or more) of boundary strength which can be identified by researchers (Price et al. 1991). It, therefore, might be the case that there is actually more information with respect to boundary strength, that is being lost in the current types of analyses.

The second issue, closely related to the first, is that information coming from boundary strength could be *against* the proposal that prosodic domains are qualitatively distinct, forming discrete prosodic domains within the structure. That is, in the current state of the intonational analysis of prosodic constituency, the notion of a *variable* boundary strength effect is implicitly discarded in favor of linking intonational events to prosodic domains. However, again, this has not been phonetically tested.

The existence of gradient differences across intonational events that are classified as distinct types of events has been present in a number of studies testing the intonation systems of languages in the past. Many areas of intonational analysis have investigated the question whether two forms are two extremes of gradient variation, or whether they belong to phonological discrete contrasts. For example, Ladd & Morton (1997) investigated whether there is a phonological difference between an 'extra' High in H*+L peak signalling contradiction (for example, 'He's iRAnian' - not Armenian) vs a neutral H*+L (for example His name is Kameiny. 'He's iRAnian'). They reported that listeners were able to distinguish the two categories, but that the result might not be because of an underlying discrete difference, but

¹Although prosodic words do not necessarily bear pitch accents; for example prosodic words following a focused item in Modern Greek. It has been shown that focus de-accent all following words in Modern Greek (Baltazani & Jun 1999).

because of the dual nature of the task. Much research has been occupied with the issue of identifying the gradient or discrete nature of tones (Ladd 1983, Pierrehumbert & Steele 1989, Gussenhoven & Rietveld 1997, among many), highlighting not only how the issue is hard to address, but also that it is a well-recognised problem within intonational analyses. Gussenhoven (1999) specifically presented an overview of techniques and some findings dealing with the question of how intonational contrasts seemingly varying in a phonetic continuum have been treated as the result of gradience or discreteness, in a paper that mainly dealt with semantically derived distinctions.

All the examples presented show how gradience has been recognised as an issue that needs to be addressed when identifying whether underlying representations of intonational events are part of a phonological distinction or not. However, gradience across boundary strength within prosodic structure was quickly discarded without investigating its phonetic manifestation, by accepting that it is irrelevant to the phonological nature of prosodic domains, which are taken to be discrete types of constituents defined on the basis of the intonational contour associated with them.

The issue of gradient variation across purported intonational events can be clearly shown to be problematic in the case of phrasal domains, like intermediate and intonational phrases. It is believed that the distinction between intermediate and intonational phrases is signalled by the differences across phrase accents and boundary tones. The definition of phrase accents vs boundary tones itself, however, heavily relies on differences of *degree* of e.g. scaling of tones, lengthening of segments etc. It is said that the phonetic differences between phrase accents and boundary tones rely on more *extreme* positions of different variables, like the occurrence of pausing, lengthening, scaling, and attached to those notions is also the notion of *similarity*. By that it is meant that many definitions of what constitutes a phrase accent (or an intermediate phrase) fall back to *comparing* it to boundary tones (or to the Intonational Phrase) and using notions like showing *more* lengthening, *more* pauses, *higher* scaling etc (Selkirk 1978, Nespor & Vogel 1986, Pierrehumbert & Beckman 1988, Arvaniti & Baltazani 2005). These definitions closely relate to the notion of *variability* across boundary types. The use of such inherently gradient properties and definitions make qualitative distinctions hard to identify, and raise

the question whether there are more boundary strength gradations than we allow to emerge by precluding that distinctions across prosodic domains are qualitative. In particular since the existence of phrase accents has not always been accepted in research, this is also questioning the existence of intermediate phrases on the basis of intonation contours. For example, Hayes & Lahiri's (1991) analysis of Bengali proposes only the levels of prosodic word and intonational phrase for the language, while Arvaniti & Baltazani's (2005) proposal for Greek also entails intermediate phrases. Phrase accents were first proposed within the autosegmental theory by Pierrehumbert (1980, as cited in Grice, Ladd & Arvaniti 2000), but there has been much research questioning their existence (Frota 1998).

To conclude, linking the definition of prosodic domains to the identification of intonational entities has mixed two different problems, the definition of intonational entities in one hand, and the definition of prosodic domains on the other. The present chapter aims at keeping the two issues distinct, and investigating support for the postulation of discrete prosodic domains on the basis of boundary strength variations.

It is, therefore, the goal of the present chapter to test whether there is acoustic support for postulating a number of discrete labelled domains on the basis of F0 marking of prosodic boundaries. Moreover, this chapter aims at investigating the nature of *variability* across boundary conditions; the boundary conditions created for the present experiment have been shown to vary gradiently from weaker to stronger. By investigating similar intonation contours across conditions, it is aimed to test whether the resulting groupings support the existence of discrete domains linking to specific intonational entities, or whether there is solely a variable output across domains, like the one found in pre-boundary durations, or possibly a combination of both effects. The present investigation is not meant to provide support for the existence of phrase accents, or for their difference to boundary tones, unless such differences can be found localised to the vicinity of the prosodic boundary. Rather, the present experiment investigates whether there is support for different types of prosodic domains on the basis of intonational marking of boundaries. Therefore, the issue of identifying abstract phonological units of intonation

events, and the one of identifying whether there are distinct labelled domains are kept distinct.

6.1.2 Modern Greek and GrToBI

Intonational marking of prosodic structure of Modern Greek has been investigated in a series of experiments performed mainly by Amalia Arvaniti and Mary Baltazani. In Arvaniti & Baltazani (2005), where they presented an adaptation of the ToBI system to Modern Greek (GrToBI), they proposed that the structure of Modern Greek entails three prosodic levels above the word: the prosodic word, intermediate phrase and intonational phrase, which, as was discussed in the Introduction Chapter, are associated with pitch accents, phrase accents and boundary tones respectively.

The present investigation aims at looking at whether there is clear acoustic evidence for supporting differences of type across conditions of varying boundary strength, as would be expected in the GrToBI view. For this reason, intonation events with similar contours at the vicinity of a variety of boundary contexts are needed. The selected constructions for the purposes of this experiment entailed LH targets before and after the boundary. In GrToBI terms, the types of intonational events that are most comparable across all the range of prosodic domains are L^*+H vs $L^* H-$ and $L^* H-H\%$. These three theoretically different intonation contours are all very common in Modern Greek; Arvaniti & Baltazani's (2005) analysis discusses how the L^*+H prenuclear accent is the most common configuration in Modern Greek. It is also said that $H-$ phrase accent usually follows an L^* pitch accent in continuation rises, which again is the common configuration in statements that comprise of at least two intermediate phrases. The use of phrases ending in L^* followed by an H target (which can be treated as a trailing tone of a pitch accent, a phrase accent or a boundary tone), allows the creation of comparable boundary conditions. Finally, Arvaniti & Baltazani (2005) state that Modern Greek can have $H-H\%$ sequences in either continuation rises, or in questioning calling contours.

Therefore, the boundary conditions used in the present experiment were designed so as to potentially entail either or all of those three distinct types of tones; BC1 is expected to be marked by L^*+H pitch accents, while conditions BC2 to BC4 can

either be marked by phrase accents H- or boundary tones H-H%, since both are possible configurations. The present experiment is aimed at testing the acoustic manifestation of the tones in the vicinity of the prosodic boundary to test whether there are differences across boundary conditions supporting the postulation of distinct phrasal domains. Before the boundary the possible configurations are always an L* followed by an H peak, which can either be the trailing tone of a pitch accent, or some type of an edge tone, and following the boundary there is always an L*+H configuration, rendering the conditions comparable.

Phonetically, there are some expectations that can be formed regarding the differences across the three types of possible tonal events based on past research. L*+H prenuclear pitch accents in Modern Greek were investigated by Arvaniti, Ladd & Mennen (1998), who reported that the Low target is consistently aligned with (or slightly before) the onset of the accented syllable, and the High peak with the first postaccentual vowel (Arvaniti & Ladd 1995, Arvaniti et al. 1998). The exact phonetic characteristics of the alignment of these prenuclear accents change when there is tonal crowding. Under tonal crowding it is possible for either the L to be undershot, or to be realized earlier than normal with an undershot H, or to be realised earlier and the H to be displaced to the right, thus aligning later than expected (Arvaniti 2007). For this reason, special care was taken in the present experiment so that the position of the stressed syllables before and after the boundary were far from each other, to allow the tonal events to reach their targets. Arvaniti (2000) reported that the “typical” alignment and scaling of the LH tones can be observed when prenuclear accents are separated by at least two unstressed syllables. In the present experiment there are four syllables intervening between the stressed syllables. Prenuclear L*+H pitch accents are expected to differ from H edge tones mainly in the H peak alignment patterns with respect to the segmental string; phrase accents are expected to primarily align with the prosodic boundary in Modern Greek (Grice, Ladd & Arvaniti 2000, Arvaniti & Baltazani 2005), as do boundary tones. Thus, the phonetic difference between an L*+H pitch accent and an L* accent followed by an H- phrase accent lies mainly in the alignment of the H target. Baltazani & Jun (1999), investigating focus intonation of Modern

Greek, tested sentences containing more than one intermediate phrases and concluded that the right edge of IP-medial ips is demarcated by an H- phrase accent, which is realised as a rise co-occurring with the last syllable of the ip.

Regarding phonetic differences between H- and H% edge tones, Arvaniti & Baltazani (2005) investigated differences in scaling between accents annotated as L* H- vs L* H-H% and reported significant differences in the scaling of the preboundary H between the two types of events, in the sense that the H peak in the ip condition was scaled lower than the IP condition. However, they reported no statistics for their experiment, making it difficult to understand whether the two proposed tones were clearly distinct, or whether their distributions overlapped in a gradient manner. Moreover, they classified the intonational events using the GrToBI annotation scheme, which might mask effects of gradience that could arise if the data were allowed to provide the groupings available, instead of the experimenter imposing the distinctions beforehand.

The existence of L and H targets before the boundary across all conditions allows for the experiment to test whether the scaling and alignment of the tones across conditions with boundaries of varying strength support the predictions offered by the GRTToBI, or whether they give rise to gradience across conditions that does not allow for the identification of specific labelled prosodic domains.

Hence, Modern Greek offers a useful setting with H preboundary targets varying both in terms of their theoretical status (part of pitch accents, or phrase accents or boundary tones), and in terms of the expectations formed for the phonetic manifestation of each target. Therefore, the main goal of the present chapter is to test the question whether there is acoustic support for proposing the existence of discrete prosodic domains based on intonation contours, or whether the issue of identifying intonational targets and prosodic domains should be kept distinct, and to conclude on whether there is support for the existence of labelled domains on the basis of their intonational marking.

6.1.3 Effect of utterance length on F0 scaling

When constructing the boundary conditions for the experiments in this thesis, the syntax and the length of the utterances were manipulated. In BC4, the phrases preceding and following the boundary were made longer than in BC3. The length of the utterances has been known to affect the scaling of the tones preceding and following the prosodic boundary, but its exact effect, and cross-linguistic differences, are issues that are still not clear (Connell 2004, del Mar Vanrell Bosch 2007). Two effects may influence the phonetic realization of tones in the vicinity of a boundary; (i) declination, that is “a phonetic effect, a backdrop lowering of F0 that has an expanded, or global, domain, often the entire utterance” (Connell 2001), and (ii) final lowering, “the progressive lowering of overall pitch range within the last 250 ms or so of an utterance” (Arvaniti & Baltazani 2005).

It is possible, therefore, that BC4 preboundary tones will be scaled lower than BC3 ones. Arvaniti (2003) investigated the effect of having more segmental material on the scaling of accents in Modern Greek, and reported that there was an effect in the anticipated direction (that is a lowering of the pre-boundary H target), but it wasn't a large effect, therefore it wasn't clear whether the declination component would affect the scaling of the accents significantly.

6.1.4 Acoustic support for labelled views

Identifying prosodic domains on the basis of intonation contours means investigating the alignment, scaling and overall shape of the contours in the vicinity of the prosodic boundary, while also taking into consideration lengthening patterns of the segments in pre-boundary positions. The alignment and scaling of tones are inherently gradient variables, that is they are suitable for expressing gradual differences varying in a continuum of possible values. Dealing with gradient variables makes it imperative to identify the types of findings that are to be taken in the present experiment as support for the postulation of domains that are of distinct type in the prosodic structure.

In the Introductory chapter to this thesis it was discussed that the main way for identifying a qualitative boundary strength effect on such inherently gradient variables is a difference in the alignment and scaling of the tones which refers to some different *linguistic* variable. Alignment patterns across boundary conditions can offer an exemplification of how different domain types can be defined. It is possible that three different boundary conditions will be accompanied by three distinct alignment patterns; that is, tones which appear to have the same contour across domains will be found to align with respect to some different linguistic variable (e.g. different syllable, stressed or unstressed, boundary etc). Such a finding would theoretically provide support for postulating qualitative distinctions across domains, not just statistical differences. Take for example the difference between alignment association of pitch accents vs boundary tones; pitch accents are known to associate with stressed syllables, and align with respect to them, while boundary tones are known to associate with respect to the prosodic boundary.

Variable differences across boundary conditions cannot provide support for the postulation of specific prosodic domains, as we discussed in the Introduction. This does not mean that there might not exist labelled domains in the structure, which are however signalled phonetically gradiently. However, if the phonetic output of those different domain types is acoustically gradient, it does not provide support for postulating different domain types. If there are underlyingly distinct domain types, they will need to be identified by a different means, e.g. a sandhi phenomenon. This thesis is concerned with investigating whether there is support for labelled views, and any gradient findings are discussed when found in the data.

6.2 Goals

The present experiment is motivated by the lack of research testing the prevalent assumption that the acoustic manifestation of intonation contours will provide support for the existence of labelled domains within the prosodic structure.

The main research questions addressed in this chapter are:

1. Is there support in the acoustic manifestation of F0 for the postulation of labelled domains? It is the main goal of this chapter to test whether boundary

strength gives rise to a limited set of distinct prosodic domains, specifically testing whether there might be reason to propose that there is variability across prosodic domains. It investigates for differences between the prosodic word and phrasal level (a difference which should be very marked, as discussed above), and for differences between levels of phrasal domains

2. How do the results from this chapter compare to those from pre-boundary lengthening?

6.3 Expectations

From the six prosodic conditions used throughout this thesis (BC0-BC5), only four can be used for the purposes of this experiment, namely BC1 to BC4. The reason is that only those conditions were created in such a manner that their intonational configurations are comparable. BC0 has the prosodic boundary within a Prosodic word, thus is excluded by definition, and BC5 has the boundary between two utterances, the final contour of which is a fall L-L%, and can thus not be compared to the H tone preceding the boundary in all other conditions. Each of those sentences are intonationally marked, and expectations can be formed regarding the type of intonational marking they will receive.

Theoretically, the boundary between the carrier words in BC1 is of the prosodic word level, and is therefore expected to be demarcated by L*+H pitch accents. The boundary between words in BC2, BC3 and BC4 is expected to be either one of intermediate phrases, therefore demarcated by H- phrase accents, or of intonational phrases, therefore demarcated by H% boundary tones. It is not clear whether both types of intonational events will emerge in the data, since it is not clear which boundary conditions will carry which types of accents. However, this is done on purpose exactly in order to test edge tones that could be of either type, and see whether there is a binary distinction emerging in the data supporting a difference between phrase accents and boundary tones.

Based on these crude expectations, it can be hypothesized that there should be an at least two-way grouping of alignment and scaling of tones; one in the shape of pitch accents, and one in the shape of phrase accents or boundary tones. Moreover, it is

possible that between conditions BC2 to BC4 there might be a qualitative difference signalling the distinction between phrase accents and boundary tones, which could either appear in the overall shape of the pre- and post-boundary contours, or in the specific alignment and scaling patterns used across purported domain. If all boundary conditions between BC2 and BC4 are of the same domain type, they are expected to be signalled in the same way.

One note here is that, if a researcher does not subscribe to the existence of phrase accents, they will have to view the edge tone of conditions BC2-BC4 as a boundary tone. In the present experiment there is no need to decide which of the two possible tones we are dealing with; irrespectively of how the tone gets transcribed, the expectation remains that the contour and alignment patterns should be the same across conditions, if they form part of the same domain type, or be qualitatively different if they form part of different domains.

In order to test whether the hypotheses just presented on the basis of the GrToBI analyses are indeed shared by phoneticians using the system, a small set of sentences entailing samples from BC1 to BC4 was sent to a trained Greek phonetician for annotation. In BC1, the preboundary word was marked as having an L*+H pitch accent, as expected. For BC2, the annotator transcribed the contour of the preboundary word as being an L* H- ending in a High phrase accent, while for conditions BC3 and BC4 the annotation was L* !H- with a downstepped phrase accent. It is, therefore, seen that in a GrToBI analysis all higher order conditions are transcribed as entailing a phrase accent. The present experiment will look for acoustic support for this annotation, as well as for acoustic support of the distinction between BC1 and the other conditions.

Regarding the possible boundary strength effects on measured variables, the following expectations can be formed:

- *Expectation A:*

The effect of boundary strength on the measured variables is qualitative, providing support for labelled views. It is expected that there must be acoustic support separating BC1 from the other conditions, and it is not clear what types of effects will arise between conditions BC2 to BC4.

- *Expectation B:*

The effect of boundary strength on the measured variables gives rise to gradient variability across boundary conditions, without providing support for any qualitative effects, therefore not clearly supporting the existence of labelled domains. This finding will need to be discussed both in terms of the difference between BC1 and the other conditions, and with respect to conditions BC2 to BC4.

- *Expectation C:*

There is a mixed result of qualitative and variable effects across conditions.

6.4 Methodology

The goal of this chapter is to test whether the alignment and scaling of F0 targets when placed under varying degrees of boundary strength show changes of qualitative nature targeting different linguistic variables across domains, supporting a view of labelled domains within the prosodic structure. Moreover, the exact phonetic behaviour of F0 targets is coupled to results from preboundary lengthening in an effort to see how the two are used together in signalling differences of type across prosodic boundaries.

6.4.1 Materials

The items used for the analysis of the intonational demarcation of prosodic constituency are taken from the datasets designed to test preboundary durations and stop voicing. This provides the experiment with a large database, and allows for direct comparisons between F0 and the durational patterns found in Chapter 3. The dataset from the experiment on the resolution of vowel hiatus was excluded because there is no clear landmark of where the boundary falls between the vowels in hiatus.

When designing the items for all the processes investigated in this thesis, the template used was the same across phenomena (see Chapter 2, Section 2.1.3). During the design, potential influencing factors on the F0 of the utterances were taken into consideration. Thus, the stress pattern for the pre- and post-boundary words is kept constant, and there are three syllables intervening between the stresses of the

pre- and post-boundary words, thus providing enough time to avoid tonal crowding, and allow for any tones to be fully realised (Arvaniti 2007). Furthermore, the elicitation technique ensured that the focus of the utterances fell at the end of each utterance, and not on the materials under investigation.

A total of approximately 640 items were analysed for this experiment (2 [data from preboundary lengthening and stop voicing experiments] * 5 speakers * 8 items * 2 repetitions * 4 boundary conditions, excluding the items that were not analysed for stop voicing and pre-boundary lengthening respectively, as described in each chapter).

6.4.2 Measurements

The test measurements were the F0 minima and maxima before and after the boundary. Using conditions BC1 to BC4, the L and H targets before (L1 and H1) and after (L2 and H2) the boundary were measured. In particular, measurements were taken for:

- the F0 of L and H before the boundary (L1 and H1)
- the F0 of L and H after the boundary (L2 and H2)
- the slopes from L1 to H1, and from H1 to L2
- the alignment of the H1 and L2 turning points with respect to the segmental end of the preboundary word

6.4.3 Contour extraction and Normalisation Technique

Two issues arise when performing F0 measurements across a large database with many speakers. First, how to automate the identification of turning points, so that all measurements are consistent, especially when the F0 contour has perturbations from stop consonants (which is the case with the dataset for this experiment). Second, how to pool across speakers whose raw F0 values carry unwanted and linguistically irrelevant variation. The first problem was addressed by using an algorithm that smooths out the F0 contours (MOMEL/INTSINT) and then identifies turning points, the second by performing a normalisation technique on the raw F0 values of all speakers, thus making them suitable for pooling. Both techniques are presented in the following sections.

6.4.3.1 Identification of turning points and smoothing of F0

The F0 measurements were extracted automatically using PRAAT (Boersma 2001), in combination with the MOMEL/INTSINT algorithm (Auran 2008, Hirst 2007). Given that the current data were not specifically designed to measure F0, but rather to measure durations and stop-voicing, manual extraction of F0 from the F0 contour presented by PRAAT would not be possible due to F0 perturbations and the existence of stop consonants. The MOMEL/INTSINT algorithm creates an approximation to the F0 contour, and then identifies turning points giving their F0 values². Each wav file was inspected, and all turning points measured by the algorithm were checked by eye, to make sure that they were not obviously wrong. The use of the algorithm was preferred over manually marking the turning points, so as to avoid the interfering factor of manual errors, and within transcriber inconsistencies, which the algorithm does not suffer from.

For the purposes of this experiment, a script was written that presented the wav file along with the spectrogram, and the turning points from the algorithm. Figure 6.2 shows the output of a finished and processed wavfile. The wav file is presented along with the spectrogram and the TextGrid file. The script opens each wavfile consecutively after the MOMEL/INTSINT algorithm has been performed, and pauses for the experimenter to check visually the result of the algorithm, and make any changes necessary. The first tier of the TextGrid contains the turning points identified by the algorithm, and the second tier provides the F0 value for each turning point from the smoothed out contour. The third tier (MOMELpoints) contains only the turning points of interest (L1, H1, L2 and H2) with their raw F0 values, which were checked by the experiments, and the final tier (segmentation) contains the segmentation of the pre- and post-boundary syllables.

²Although there is a theoretical background for which the particular algorithm was built, in this analysis no such relationship is assumed. The algorithm is only used for its predictive power over where turning points fall, and for its potential of smoothing out F0 contours. Thus, the names of the turning points and their phonological status is not a matter of this debate. The MOMEL part of the algorithm, using a micro- and macro-melodic component approximates the F0 contour, while the INTSINT part converts the output of the MOMEL to a sequence of discrete tonal symbols (see Hirst & Espesser 1993, for an analysis of the micro- and macro-melodic components). These symbols are not used in our analysis, but are turned into Lows and Highs.

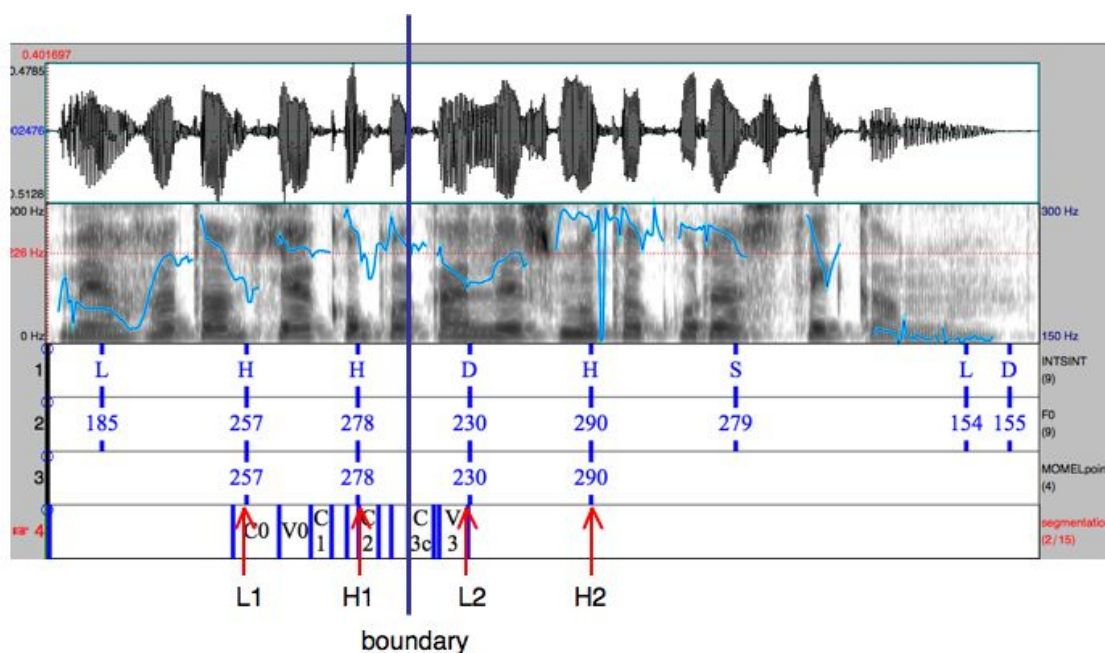


Figure 6.2: Example of identification of L and H turning points. Tiers 1 and 2 contain information that comes automatically from the MOMEL/INTSINT algorithm. Tier 1 shows the identified turning points, and Tier 2 gives their values from the smoothed out F0 contour. Tier 3 has only the four turning points under investigation, which have been copied by the experimenter. Tier 4 contains the segmentation of the pre- and post-boundary syllables. The depicted item is [a'ðjakopa ka'lezmata], and the example is taken from BC1.

6.4.3.2 Normalisation across speakers

The previous section explained how F0 values were extracted from the wavfiles in a principled and automatic manner. Following is the analysis of how raw F0 values were normalised for statistical pooling across speakers.

Each speaker has a different F0 range, with some speakers belonging higher in the F0 scale, and others lower. The idiosyncratic differences across speakers make it impossible to use their raw F0 data and pool them across speakers for the statistical analyses. Given, however, that statistical power is very important, pooling across speakers is necessary. For that reason, a normalisation technique is used, as described in Calhoun (2006, pp. 158-9).

The goal is to transform the range of each speaker into a scale from 0-100%, where 0% is the lowest point of the speaker's range, and 100% the highest (see Figure 6.3

for a schematic explanation of the algorithm). A log transformation is used (log Hz), which is meant to correspond to ratios in Hz (see Traunmüller 1990, as cited in Moreton & Thomas 2007).

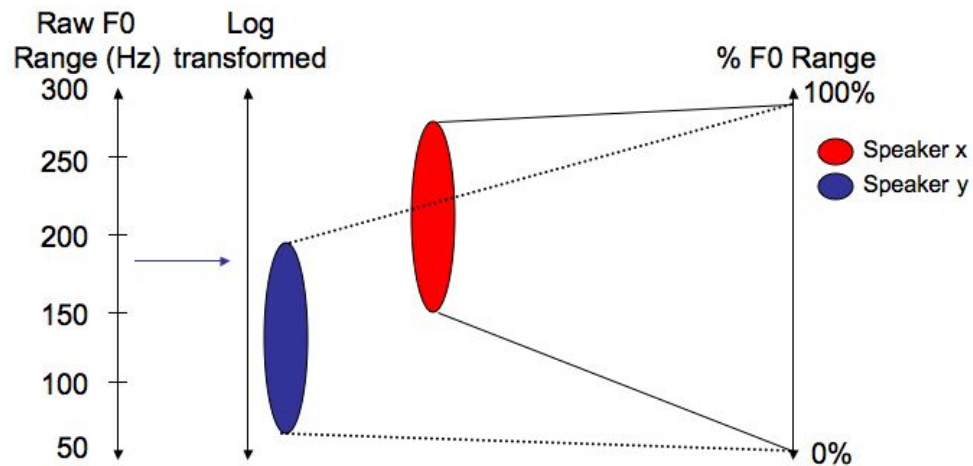


Figure 6.3: Schematic representation of normalisation algorithm. In the algorithm the raw F0 values are first computed to logarithmic, as shown in the left hand side on the y axis, and then transformed to percent scale.

First, measurements from all the productions of each speaker were taken; F0 measurements were made every 10ms in each of the sentences produced. Each speaker had approximately 6 minutes of speech, resulting in approximately 36,000 measurements per speaker. The extracted values were sorted (from lowest to highest), log-transformed, and outliers were removed by excluding values more than 2.5 standard deviations away from each speaker’s mean. The goal of this technique is to convert all F0 values from each speaker to something representing their *range* instead of the raw fundamental frequency value. Thus, taking measurements from all of their utterances assures that the range of each speaker is identified.

The next step is to convert the logged values to a scale from 0-100%. Assuming, for example, there are 35,000 values from one speaker representing the whole of their F0 range, the first 35 values are the 0.1% of their range, the second 35 are the 0.2% and so on and so forth. This way there are 1000 bins containing 35 values each, and each of these bins represents the 0.1% of that speaker’s range. This gives a database with the range of each speaker in Hz and log-transformed values. Once this database is created, and the 1000 bins established, the values extracted by the

script for L1, H1, L2 and H2 (which have also been log-transformed) are taken in turn, and using a "look-up" function they are placed in the position of the range of that speaker where each corresponds. For example, one value of H2 might fall in the 998th bin, thus belonging in the 99.8% of that speaker's range. Thus, each value now represents a percentage, and all the values are able to be pooled together and used for statistical purposes.

6.4.3.3 Alignment

Since F0 turning points were identified by the MOMEL algorithm, the position of the targets with respect to the boundary also relies on the identification of the turning points. The segmentation of the wavefile was performed by the experimenter. Therefore, for the pre-boundary High we subtracted the position of the identified H1 point from the position of the boundary; e.g. if the boundary was at 1100ms within the waveform, and the position of the H1 was at 1000ms, then the H1 is aligned at 100ms before the boundary. Similarly for L2 we subtracted the position of the L2 from the boundary (since it follows it), therefore, if L2 is at 1150ms into the waveform, then it is aligned 50ms after the boundary.

6.5 Results

Four REPEATED MEASURES ANOVAs were used to test the effect of boundary strength on the scaling of the proportional F0 targets, i.e. L1, H1, L2 and H2, with boundary strength as the independent variable having four levels (BC1-BC4). The alignment of the H tone with respect to the boundary position (the end of the preboundary word) was again tested using a REPEATED MEASURES ANOVA with boundary strength as the independent variable.

6.5.1 Scaling

The effect of BOUNDARY STRENGTH was significant for the scaling of L1, H1 and L2, and not for H2 [L1: $F(3, 360) = 180.9, p < .001$, H1: $F(3, 318) = 13.9, p < .001$, L2: $F(3, 315) = 26.5, p < .001$, H2: $F < 1$]. Figure 6.4 shows the shape of the intonational

contour for each boundary condition (duration not represented in this graph)³. Table 6.1 shows the mean values for all boundary conditions for all dependent variables (F0 values transformed as described above). Alignment measurements are in milliseconds showing the distance of the H peak to the boundary (i.e. the end of the preboundary word). The last column shows the Bonferroni post-hoc analysis, with arrows pointing to differences across conditions that were found to be significant (greater than, or smaller than the mentioned condition).

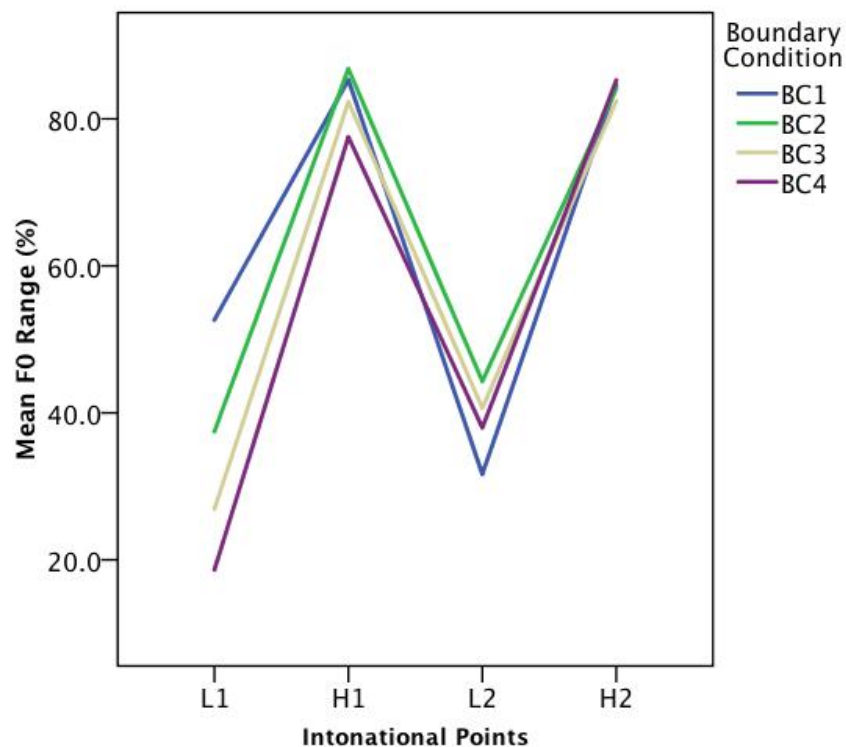


Figure 6.4: Mean range values pooled across speakers for each measurement point. Duration is not represented in this graph.

Looking at each intonational target individually, the following observations can be made. First regarding the preboundary tones, the scaling of the L1 target assumes the highest position in BC1 (54% of the speakers' range) and the lowest in BC4 (19% of the speakers' range) with in-between values for BC2 and BC3, resulting in variable differences across conditions. Similarly, the scaling of the H1 tone ranges from assuming the highest position in BC1 and BC2, followed by BC3 and BC4

³The same pattern of intonational contours can be found across speakers - see Figure B.1 in Appendix B.

		BOUNDARY STRENGTH EFFECT ON F0 RANGE VALUES				
		BC1	BC2	BC3	BC4	Post-Hoc Bonf.
Scaling (% range)	L1	54	38	27	19	1>2>3>4
	H1	85	87	83	77	4<123,2<3
	L2	32	45	41	38	1<2>34
	H2	85	84	83	84.7	-
Alignment (ms)	H1	97	43	41	33	1>234
	L2	141	138	126	133	-
Slope (% difference)	H1toL2	53	42	42	40	1<234
	L1 to H1	32	50	55	60	1<2<3<4

Table 6.1: Results on the significance of prosodic boundary strength on F0 measurement points (L1, H1, L2 and H2).

with gradually lower values. Subsequently, the slope from L1 to H1 is influenced gradiently by boundary strength (BC1 having the shortest slope, and BC4 the longest).

In the post-boundary word, the scaling of the L2 target was found to be influenced categorically by boundary strength in that BC1 assumes the lowest position, while there are gradiently varying positions between BC2 and BC4 ranging from higher to lower. Therefore, a binary distinction is clear between BC1 and BC2-BC4 L2 scaling. H2, on the other hand, is not affected by boundary strength at all. Therefore, it is not the whole pitch accent that seems to be affected by boundary strength, as would be expected if there was preplanning for longer utterances (with e.g. BC4 having lower values because of the utterance following the boundary being longer). Only the Low target of the post-boundary pitch accent is affected, suggesting that the differences between BC2 to BC4 might be an immediate effect of boundary strength, while the difference in BC1 is actually one of different intonational event, and potentially qualitatively different boundary.

A categorical difference between BC1 and the other conditions is also clear when looking at the whole intonation contour before and after the boundary. Figure 6.4 shows the mean contour across speakers for all boundary conditions, and it is clearly shown that BC1 is qualitatively different to the other conditions by having a markedly different overall shape (highest L1, lowest L2, longest H1 to L2 excursion).

Across conditions BC2 and BC4, however, the changes for all F0 targets are gradient in the sense that they vary on a continuum. Both in pre- and in post-boundary conditions the scaling of the L targets is gradually lower from BC2 to BC4. First, this pattern initially supports Expectation A in predicting a categorical difference between BC1 and the other conditions, but also gives rise to a general view of gradient in scaling across conditions BC2 to BC4. It is by itself interesting to note that the scaling of the turning points could be viewed as an added means of signalling qualitatively distinct domains, and it is also important to note that the distinction actually comes from the post-boundary rather than the pre-boundary tones. Second, regarding the gradient found across BC2 to BC4, it is shown that all three conditions must form part of the same prosodic domain, since the difference between H- and H% would be in that H% is scaled *higher* than H-, but on the contrary in the present boundary conditions the preboundary H is scaled lower in higher conditions.

If it is agreed that conditions BC2 to BC4 are of the same domain type, then it is important to explain the variable results found within this domain. One solution is the one offered by the GrToBI transcriber, that is that BC2 ends in an L* H-sequence, while conditions BC3 and BC4 end in L* !H- sequences. However, while this can explain the statistical difference between BC2 vs BC3 and BC4, it does not explain the one between BC3 and BC4. Moreover, it is clear that the effect is of a gradient nature, representing the influence of boundary strength, especially when looking at the post-boundary L2 target, which is also gradiently influenced. In addition, the differences between BC2 and BC3 is the syntactic construction, while between BC3 and BC4 it is the length of the utterances that has been manipulated. However, the acoustic differences across such different manipulations are similar, in lowering the scaling of the targets. These findings are taken to represent variable differences of boundary strength, which are present in the prosodic structure, but which the current version of labelled views is not able to express.

To summarize, the data from F0 scaling shows that the existence of distinct domain types is supported, in that BC1 is categorically different to the other domains, and in that the higher conditions all seem to group in one other type of domain. However, there are also variable outputs across conditions, which go against the analysis

offered by labelled views. Labelled views expect either the same contours across conditions if they are part of the same domain, or categorically distinct contours if their conditions belong to different prosodic domains. The finding of consistently variable outputs across domains supports a view of gradience within the structure. This finding is very similar to Ladd's (1988), supporting Expectation C, and will be further discussed in Chapter 7.

6.5.2 Alignment

The alignment of the H1 and L2 tones with respect to the segmental string was tested, with the goal of identifying qualitative differences between domain types, and in particular -if there are any- whether they reflect the expectation that the H tone for BC1 (or else the Prosodic Word) will align differently to BC2-BC4 (one is expected to align with respect to the stressed syllable, and BC2-4 with respect to the boundary).

The effect of boundary strength on the alignment of H1 with respect to the boundary was found to be significant ($F(3, 249) = 34.1, p < .001$). Bonferroni post-hoc tests (presented in Table 6.1) show that BC1 is statistically different to the other domains in its alignment pattern, and there are no differences across conditions BC2-BC4. Importantly, H1 peak in BC1 seems to associate with a different linguistic variable, giving rise to a qualitative difference between the two groups; H1 peak falls on the penultimate stressed syllable for BC1, while for conditions BC2 to BC4 it falls within the ultimate syllable, close to the boundary. This links well with the expectation that the H target in BC1 will associate with the first post-accentual vowel, while H1 peaks in conditions BC2-BC4 will all associate with the pre-boundary vowel, and while the present experiment does not test this exact issue of association, it shows that there is a clear distinction in the alignment between the BC1 on the one hand, and all other conditions on the other. Figure 6.5 shows the qualitative distinction in terms of the linguistic variable in which the H peak falls.

This finding is in accord with Expectation A, and with Arvaniti et al.'s (1998) finding regarding prenuclear pitch accents in Greek, in the sense that the H1 aligns with respect to the stressed syllable in BC1, and its alignment falls within the first post-accentual vowel. Arvaniti et al. (1998) also tested words with antepenultimate

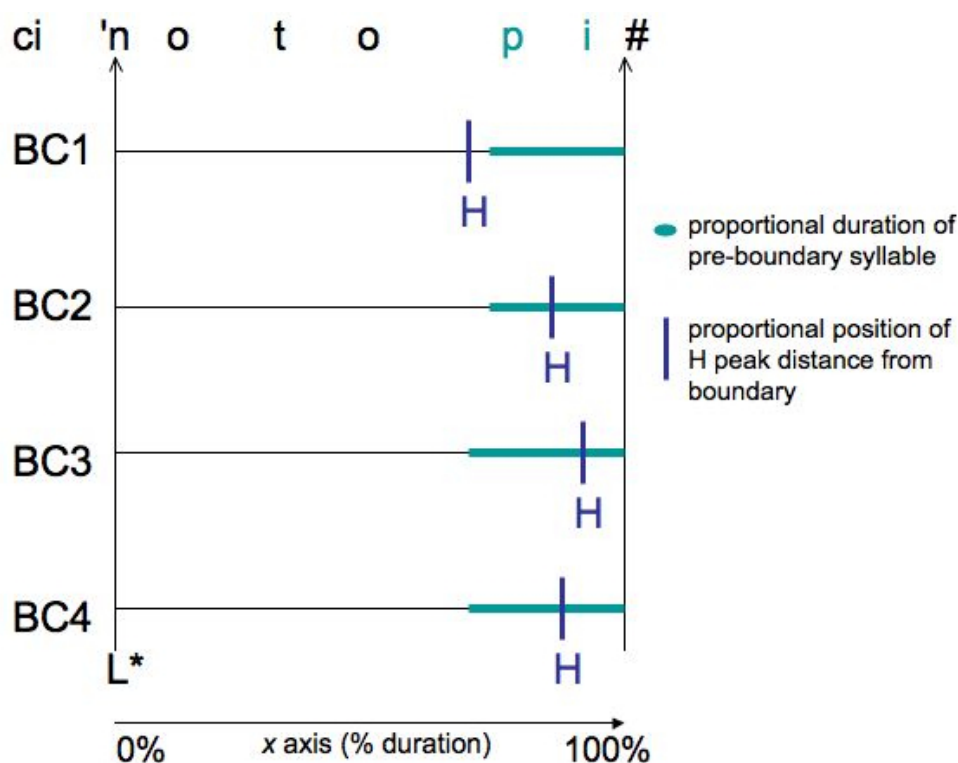


Figure 6.5: Crude schematic representation of H alignment within the pre-boundary word. All durations are presented proportionally; i.e. the distance from L to H is the 100%, and the duration of the preboundary syllable and alignment of H1 are expressed proportionally to that. For reference reasons the duration of the pre-boundary syllable is also presented (the green line); the alignment of H1 in conditions BC2-BC4 clearly falls within the pre-boundary syllable.

stress, thus their findings are directly comparable to this experiment. Moreover, it is clear that the preboundary High in all other conditions is an edge tone, irrespective of whether it is a phrase accent or boundary tone.

The alignment of L2 was also tested, looking for possible effects of boundary strength. In general, the alignment of the L* is not expected to be affected by boundary strength; its behaviour is known to be fairly stable, except in conditions of tonal crowding. Boundary strength, as expected, was found not to exercise a significant effect on this variable (Table 6.1).

Combining the effects of boundary strength on scaling and alignment, it can be concluded that there is a categorical distinction in scaling of L2 and a qualitative

distinction in alignment of H1 between BC1 and the other conditions, supporting the existence of labelled domains signalled by means of prosodic structure, but there is also support for a view of gradient differences needing to be allowed within the prosodic structure, as seen by the gradually varying values in the scaling of L1, H1 and L2 tones in conditions BC2 to BC4. The constant alignment of H1 with respect to the boundary in conditions BC2 to BC4 suggests that the H targets viewed are indeed edge tones, and that they do not provide support for a categorical distinction across conditions, but they do support the existence of gradient within the structure, as Ladd reported. Both syntactic and length manipulations had the same effect on those higher domains, and suggest that both are structural properties that need to be accounted for when defining prosodic structure.

6.5.3 Scaling and pre-boundary durations

One final issue to be discussed is the combined effect of F0 scaling and pre-boundary lengthening in demarcating differences between phrase accents and boundary tones. The expectation from the intonational approach is that the acoustic difference across phrase accents and boundary tones in the case of H- vs H% configurations is that they will differ in terms of their relative scaling and pre-boundary lengthening. In order for the distinction to arise, it means that across a set of variable boundary strength conditions a binary distinction will arise supporting the difference between phrase accents and boundary tones. However, the results from both the scaling of targets between BC2 and BC4, and the results of pre-boundary lengthening across conditions BC2 to BC4, as seen in Chapter 2, showed that the effect was gradient with gradiently varying scaling patterns, and gradiently longer durations across conditions. Therefore, it is not possible to propose a binary distinction on the basis of either of the two cues, contrary to the expectations formed by GrToBI.

6.6 Discussion

This chapter tested the effect of boundary strength on the alignment and scaling of F0 pre- and post-boundary targets in Modern Greek. The goal was to see whether labelled domains could be identified on the basis of the acoustic manifestation of F0

targets, and to compare this effect to the one exercised on pre- and post-boundary durations.

The first result that emerged from the data was that there is support for the postulation of labelled domains; a distinction was found between condition BC1, i.e. a boundary across prosodic words, and all other conditions (BC2 to BC4). There was a two-way distinction in the alignment pattern of the pre-boundary H target, in the sense that in BC1 it aligned with the first postaccentual vowel, as expected in a bitonal L*+H pitch accent associated with a prosodic word level, while higher order domains aligned with respect to the prosodic boundary. Moreover, it was found that the overall shape of the scaling of the intonational contour before and after the boundary was markedly different in BC1 than in all other conditions, providing clear support for a qualitative distinction between the word and phrasal domains. This qualitative distinction is in accord with Expectation A.

Between conditions BC2 to BC4, i.e. the ones above the prosodic word level, no categorical differences were found either in alignment or in scaling of targets. These three boundary conditions grouped together in terms of their alignment pattern, supporting the view that they are all edge tones, since they align with respect to the prosodic boundary, and had no statistical differences between them. However, the scaling of the preboundary L and H targets were found to be gradiently influenced by boundary strength. This is interpreted as showing that conditions BC2 to BC4 represent a phrase domain of some sort, which is markedly different to the lower domain, while the consistent nature of the gradient patterns across speakers is interpreted as mirroring *strength* across boundary conditions. Therefore, it is taken that all three conditions (BC2 to BC4) are different strength manifestations of the same domain. Irrespective of whether this domain is to be viewed as an intermediate phrase or intonational phrase domain, and whether the tonal events associated with it are to be transcribed as phrase accents or boundary tones, it is shown that the existence of phonetic gradience within that domain is present, and that there is no acoustic support for postulating the existence of separate tonal events, or separate prosodic domains.

This interpretation of the findings supports Expectation B, where it was predicted that the H targets would range in a continuum of possible values without providing support for distinct labelled domains. However, this finding in combination with the clear categorical and qualitative distinction between BC1 and higher order conditions clearly provides support for Expectation C, where both types of events are predicted, in accord with Ladd's analysis of the possible nature of prosodic constituency.

It therefore seems that, since the phonetic manifestation of the different levels mirrors differences of strength, all three conditions must come from the same type of prosodic domain. In a GrToBI analysis, BC2 final phrase accents would be transcribed as H-, while BC3 and BC4 would be transcribed as !H-, i.e. the downstepped High tone, as was done by the expert phonetician. However, annotating them all as phrase accents means that it is accepted that they are associated with intermediate phrases, which are all of the same strength. If they are of the same prosodic domain, there is no way to explain the difference in scaling between conditions BC3 and BC4, since they are both transcribed as !H-, but are phonetically realised with different scaling. If, on the other hand, the differences across conditions were to be attributed to final lowering, then there is no way to explain the difference between conditions BC2 and BC3; final lowering is expected to be more in longer sentences, but conditions BC2 and BC3 do not differ in terms of their length, therefore should not differ in terms of the amount of final lowering. By this it is meant that acoustically there was a gradient change across BC2 to BC4, with differences across all conditions, but in a GrToBI analysis only a two way distinction between downstepped and non-downstepped tones exists.

These findings highlight how the available tones for capturing differences across utterances are not enough to mirror all possible strength differences that are produced by speakers of Modern Greek. Importantly, however, they show that postulating another tone, or different levels of downstepping, to account for them again is not the proper way of dealing with the matter, since there is no clear acoustic property to separate categorically one from the other. Rather, they seem to vary on a continuum of possible strengths, which are derived from syntactic and length manipulations.

Importantly, the present results show that a syntactic manipulation (i.e. the difference between the constructions BC2 and BC3) and a length manipulation (i.e. the difference between constructions BC3 and BC4) have both the same type of acoustic effect on the preboundary H target: the scaling of the H peak gets gradually scaled lower. This finding is consistent across speakers and across phenomena; the duration of pre-boundary segments has been shown to get gradually longer between conditions BC2 to BC4, tone alignment was shown to vary gradiently, and less coarticulation has been shown to occur in vowel hiatus resolution across BC2 to BC4. Therefore, the effect that a large syntactic manipulation exercises on the production of segments is similar to that exercised by the length manipulation, placing the two in equal positions in the prosodic structure. This finding has two implications; first, that such effects can represent differences of *strength*, not type as the labelled views predict, and second, that these effects both need to be incorporated in the prosodic structure schema in a way that acknowledges the possible effects from both syntax and length (this issue will be discussed in Chapter 7).

One issue to be noted is that the alignment of the H targets seems to be less prone to gradient influences from boundary strength in the data, while scaling allows for more variation. The boundary strength effect on the two types of measurements was manifested differently in conditions BC2 to BC4; alignment of the H target did not seem to be influenced at all, while scaling was influenced gradiently giving rise to the main differences across those three domains. It could be proposed that the alignment of the H tone specifies the nature of the tone as edge tone, while the scaling mirrors the differences in boundary strength, therefore highlights the need for having a more elaborate representation of prosodic structure. However, the distinction between the role of the two variables was not the matter of investigation of the present experiment, so no clear conclusions can be drawn other than the observation.

The findings in the present chapter link very well with the ones in the previous chapters; results from post-nasal stop voicing supported the existence of a labelled domain, like the prosodic word, a finding that is replicated in the present chapter. Results from pre-boundary lengthening and vowel hiatus on the other hand provided initial support for the postulation of gradience within the structure; syntactic and

length manipulations both seem to give rise to gradient differences across conditions BC2 to BC5, a finding that was again clearly replicated in this chapter. This prevalent finding across experiments and speakers suggests that we are dealing with one type of domain across BC2 to BC4 varying solely in terms of relative strength across boundary contexts, which are however signalled in the acoustic output, and look like they should be represented in the prosodic structure. These results will be analyzed further in the following and final chapter of this thesis, where we will discuss what they suggest for the nature of prosodic constituency, and what their implications are for current labelled views on prosodic structure.

CHAPTER 7

Discussion

This concludes the examination of the acoustic manifestation of boundary strength in Modern Greek, which set out to investigate whether acoustic correlates to prosodic structure support the existence of labelled domains. We end this thesis by summarizing the main theoretical claims of this work and the main findings of each chapter, along with a discussion of the ramifications of the theoretical claims for past research within the labelled views.

7.1 General Summary

The experimental investigation of the effect of boundary strength on the acoustic manifestation of segmental and suprasegmental processes in this thesis has provided evidence for a number of claims. Some of them support past research on prosodic constituency, and some add findings to be taken into consideration.

First of all, evidence from three of the four phenomena investigated provide support for the hierarchical nature of prosodic structure. Pre-boundary syllables were gradually longer in higher boundary conditions, as has been found in many languages in the past (e.g. Wightman et al. 1992, Cambier-Langeveld 1997). Moreover, support was provided by patterns of coarticulation. Vowel pairs in hiatus were found to be resolved using coarticulation, and the effect of boundary strength on formant values and duration patterns revealed more coarticulation in lower domains (similar findings can be found in Tserdanelis 2004). That is, somewhat more extreme formant positions were found across stronger domain boundaries, suggesting more ‘typical’

values, along with longer durations. Finally, F0 scaling at the phrase level showed a variable effect supporting a hierarchical structure, with lower pre-boundary H peak scaling in higher boundary strength conditions.

The main question we set out to investigate was whether there is acoustic support for the postulation of labelled domains. Two out of four phenomena investigated were found to be influenced qualitatively by boundary strength, supporting the existence of labelled domains. Post-nasal stop voicing was found to take place optionally within the lowest boundary condition investigated, i.e. a particle and a verb, and to be categorically blocked across all higher conditions, providing support for a qualitative boundary strength effect, as anticipated by labelled views. Similarly, F0 alignment and scaling results, as was anticipated, distinguished the word from the phrasal domains in a categorical fashion, supporting the acoustic difference between pitch accents and edge tones, hence also supporting the distinction between lower and higher prosodic domains.

Parallel to the evidence in support of labelled domains, this thesis also showed variable boundary strength effects that cannot be accounted for by labelled domains. Some effects of boundary strength gave rise to consistent differences across boundary conditions which can only be represented in the prosodic structure if degrees of boundary strength within the same domain type are allowed. This issue is further discussed in Section 7.4, where possible options for combining labelled domains with recursivity are discussed.

Therefore, the present thesis has provided reliable acoustic evidence for the concurrent existence of both differences of *type* and differences of *strength* across boundary conditions.

The final main observation from the findings of this thesis is that the prosodic grouping of utterances was shown to be influenced by the length of the utterances involved. It was shown that longer utterances gave rise to stronger boundary strength than shorter utterances. Acoustic evidence supporting this finding came from pre-boundary lengthening, vowel hiatus resolution and intonational marking of prosodic phrases. Pre-boundary durations were found to be longer in the vicinity

of stronger constituents, vowel hiatus resolution took place with less coarticulation, and there was gradually lower scaling of F0 across conditions with stronger boundaries.

Manipulating the length of the utterances was shown to give rise to differences of strength across conditions: longer utterances only gave rise to variable differences, not categorical ones. Importantly, the effect was found to be similar to manipulating the syntactic construction of the relevant utterances. The difference in boundary strength across conditions whose syntax was changed was comparable to the difference in boundary strength across conditions whose length was longer. Moreover, neither the syntactic nor the length manipulations gave rise to categorical effects.

All these findings are discussed with respect to past research in the rest of this chapter.

7.2 Methodological contribution

The present thesis has used an innovative approach to investigating boundary strength effects in the sense that it investigated acoustically a variety of phenomena in parallel in one language using instrumental and perceptual analyses. The concurrent use of several processes has been very important in providing solid evidence both for the existence of labelled domains, and for the possibility of differences of boundary strength within the same domain type within the structure. Pre-boundary durations proved how the experimental design managed to capture a variety of boundary strength conditions, ranging from weak to strong. The magnitude of lengthening from one condition to the others was also shown to be similar; for example, the lengthening between BC1 to BC2 was similar as that from BC2 to BC3. This gradual lengthening showed how the conditions investigated captured the variability of possible boundary strengths. Therefore any categorical effects found across domains in the manifestation of the other phenomena are reliably thought to mirror qualitatively different domains of the prosodic structure.

Hence, the investigation of pre-boundary lengthening concurrently to the other domains was of crucial importance for the thesis. However, the other three phenomena were also of significance with respect to the findings, since each contributed in a

different way in our understanding of prosodic constituency. The two processes traditionally viewed as sandhi-type phenomena allowed insights into how and of what type boundary strength effects can be identified when using sandhi and coarticulatory processes, while F0 alignment and scaling tackled the issue of intonational marking of constituents, one of the most well-known and used correlates for identifying labelled prosodic domains.

To sum up, the methodology used in this thesis makes it one of the very few studies to have investigated a variety of phenomena with respect to their delimitative function for prosodic constituency, and therefore provides the thesis with an overview of possible effects and findings that most studies do not have. It is often very difficult for a researcher to compare across studies, languages and experimental conditions when trying to understand prosodic constituency. This thesis takes an important step towards understanding the interplay of so many acoustic correlates and the prosodic structure, and towards understanding the types of effects that can be viewed as mirroring differences of type vs differences of strength across domains.

7.3 Summary of chapters

Four chapters were devoted to the investigation of boundary strength effects on the acoustic manifestation of four phenomena respectively. Chapter 3 investigated pre- and post-boundary durations, Chapters 4 and 5 investigated two sandhi-type phenomena, namely post-nasal stop voicing and vowel hiatus resolution, and Chapter 6 looked at the effect of boundary strength on the alignment and scaling of F0 targets before and after the boundary.

Chapter 3 showed that Modern Greek has final lengthening above the word level in pre-boundary positions. Stronger boundaries were found to give rise to longer pre-boundary durations. The whole pre-boundary syllable was found to be lengthened by the boundary effect. The effect was also apparent on the penultimate and antepenultimate syllables. The amount of lengthening from the lowest to the higher condition was shown to be more for the *antepenultimate* syllable, in comparison to the penultimate one. This is in support to the proposal that linguistically ‘important’ syllables might in some way attract the lengthening effect, since

the antepenultimate syllable is stressed (Turk & Shattuck-Hufnagel 2007, Turk & Dimitrova 2007).

This chapter also provided insights into boundary strength effects on the duration of post-boundary segments. This effect had not been previously investigated for Modern Greek. It was shown that Modern Greek does not exhibit post-boundary lengthening. A small non-statistically significant tendency was sometimes found on the closure of the post-boundary stop showing some lengthening in higher domains, but the effect is not significant. This shows that pre-boundary lengthening is the main durational marker of prosodic structure in Modern Greek.

Importantly, this chapter provided solid evidence for the experimental design of this thesis; the pattern of gradual lengthening across domains showed that the boundary conditions created varied in terms of their relative strength.

Chapter 4 on post-nasal stop voicing provides the first piece of evidence in support of labelled views. It is shown that the perception and production of post-nasal stop voicing in Modern Greek distinguishes two qualitatively different domains. The application of the phenomenon is allowed across the lowest of the investigated conditions, i.e. between a particle and a verb, and it blocked across all other higher conditions.

One interesting observation with respect to post-nasal stop voicing occurrence is that the experiment revealed differences of patterns across lexical items. That is, some lexical items allowed voicing to occur, while others didn't. For example, while voicing occurred 100% between the negation particle [ðen] and a verb, it was blocked across the temporal and subjunctive particles ['otan] and [an]. At the same time, voicing took place approximately half of the times for the other negation particle used ([min]). This finding partially supported Nespor & Vogel's (1986) analysis, according to which this difference in lexical items was used to highlight the distinction between a phonological word and a Clitic group. According to them stop voicing is optionally allowed across phonological words, and blocked across clitic groups. Therefore, they propose that the difference in prosodic structure is expected to be:

ðεν παίζω '(I) don't play' → [_C [ω ðen]ω [ω 'pezo]ω]_C → ðe'bezo
 αν παίζω 'if (I) play' → [_C [ω an]ω]_C [_C [ω 'pezo]ω]_C → *abezo

However, it is not clear why there is this difference across lexical items. Moreover, the application of stop voicing for the negation particle [ðen] seems to be an obligatory rule, since it takes place 100% of the time. Under Nespor & Vogel's (1986) analysis, stop voicing is only obligatory within phonological words. Therefore, the analysis should really be:

ðεν παίζω [_C [ω ðen 'pezo]ω]_C → ðe'bezo

But that goes against their definition of a phonological word, since the particle needs to attach as a clitic to the verb host and form the clitic group. Therefore, while their analysis captured the difference in lexical items, it was not able to explain all differences that can be identified on the basis of instrumental data.

Additionally, it was found that there was a number of items which, when played back to the raters for the perceptual classification of whether voicing had taken place or not, the raters did not reach agreement. The phonetic realization of those instances more closely resembled voiceless stops than voiced, but the lack of agreement suggested that maybe there are some patterns of coarticulation for some of the items, suggesting that within this lowest domain not only is the rule optional, but also that maybe the phenomenon could be classified as the result of coarticulation. A different explanation could be that voicing some lexical items has become phonological, but not for some others. However, all those issues should constitute the matter of future research, since the present experiment did not set out to investigate them.

Hence, while this experiment has provided partial support for Nespor & Vogel's (1986) impressionistic analysis in providing evidence for a qualitative distinction across conditions which allow or block the application of stop voicing, and by identifying the difference across lexical items they had found. The experimental data also revealed that Nespor and Vogel's analysis of stop voicing proposing the existence of a Clitic group did not provide some type of identifiable measure for knowing when stop voicing is allowed to occur and when it isn't. This gives rise to a circular reasoning of identifying a prosodic domain on the basis of a phenomenon, and identifying where a phenomenon will occur on the basis of the domain across which it is expected to be blocked or allowed.

The investigation of sandhi-type phenomena and its ability to signal differences of domain type was further analyzed in Chapter 5 by looking at vowel hiatus resolution in Modern Greek. Although the phenomenon has been reported to be coarticulatory (Baltazani 2006, Kainada 2007), the effect of boundary strength on it might still be categorical. Therefore, even if traditional impressionistic analyses by Nespor & Vogel (1986) and Condoravdi (1990) treated it as a sandhi phenomenon, it could still be the case that it signals labelled prosodic domains, as was suggested by more recent acoustic research by Pelekanou & Arvaniti (2001) and Baltazani (2006). These last two studies, in particular, had given rise to a mismatch in the domain across which they proposed that coarticulation was blocked, providing further motivation for an instrumental investigation of the phenomenon.

Results showed that there was gradually less coarticulation across stronger boundary conditions, measured in the behaviour of formant values and duration patterns. Therefore, contrary to previous research, this chapter clearly showed that vowel hiatus is resolved with coarticulation, at least in the domains investigated, and that it does not give rise to qualitative differences across investigated domains.

Importantly, the duration of the vowel pairs was not found to correlate with the position of formant values, suggesting that the two properties of the production of vowel hiatus are distinct, and that it isn't necessarily the case that formant values assume their 'typical' positions when there is time for them to do so. This finding was contrary to Baltazani's (2006) study, and needs further testing with articulatory measurements in order to investigate the mechanisms regulating vowel hiatus resolution in Modern Greek.

The final correlate to prosodic constituency investigated in this thesis was the alignment and scaling of F0 targets before and after the boundary, which was analysed in Chapter 6. We investigated conditions BC1 to BC4 with the expectation that BC1 would be different to the other, since BC1 constructions were expected to associate with pitch accents, while BC2 to BC4 with edge tones. Using similar intonation targets we showed that the anticipated qualitative distinction did arise between BC1 and the other conditions. Additionally, it was shown that no qualitative distinction arose across conditions BC2 to BC4, and that there were consistent

differences of boundary strength across all three conditions. This supports a view of a mixture of differences of type and strength within the prosodic structure, which will be discussed in the following section.

7.4 Labelled domains and recursivity

The main research question this thesis set out to investigate was whether there is support in the acoustic manifestation of boundary strength for the postulation of labelled domains within the prosodic structure, or whether the acoustic manifestation of prosodic boundaries only give rise to differences of boundary strength. The thesis showed results that support the concurrent existence of both labelled domains and variable differences of strength within domain types.

Support for labelled domains comes, as discussed above, from the application of post-nasal stop voicing and the intonational marking of constituency. On the contrary, vowel hiatus resolution proved to be affected variably by boundary strength, contrary to previous analyses of the phenomenon by Nespor & Vogel (1986) and Pelekanou & Arvaniti (2001). This result showed that traditional impressionistic analyses of phenomena do need to be investigated instrumentally, since it is possible that conclusions proposing the existence of distinct domain types on the basis of such phenomena can be based on the lack of instrumental data.

On the basis of the present data we form two hypotheses. First, it seems that there is a qualitative distinction between lower and higher domains in the prosodic structure. Lower domains seem to entail and be marked by phenomena that apply categorically (post-nasal stop voicing and F0 differences stemming from rhythmic differences), and not to have gradations of boundary strength, while higher conditions seem to be produced in a variety of boundary strength levels, and are mainly marked by inherently gradient phenomena. Second, in the present data it seems that the phenomena signalling categorical distinctions are phonological. Post-nasal stop voicing seems to be a sandhi rule. F0 alignment distinguished two phonologically distinct parts of the structure, i.e. the word from the phrasal level, which are also distinguished on the basis of their rhythmic structure. On the other hand, pre-boundary lengthening, F0 scaling within phrasal domains, and vowel hiatus

resolution by coarticulation all undergo a variable boundary strength effect which does not give rise to differences of a qualitative nature. Hence, on the basis of the present data it could be hypothesized that phonological phenomena are more prone to signal labelled distinctions, while inherently gradient variables are more prone to mirror variable boundary strength effects, and that this distinction also separates lower from higher domains.

However, the above stated hypotheses are strong claims to be tested looking at literature across phenomena from cross-linguistic research. The important part of the proposal is that there seems to be some qualitative difference between lower and higher domains with lower domains seemingly attracting more phonological phenomena than higher ones. This could give rise to a binary distinction within the prosodic structure, and could explain how many of the sandhi phenomena used to postulate the existence of phrasal domains have been shown by instrumental analyses to be influenced variably by boundary strength.

Support for the existence of variability in the prosodic structure comes mainly from F0 marking of phrasal domains. It was seen that, while the distinction between the word and phrasal levels was clearly demarcated on the acoustic signal both in the alignment of the H peak, and in the overall intonation contour, distinctions across domains of phrasal level were varying gradiently. This finding might suggest that either there are no categorical differences across those conditions, or that none can be identified on the basis of this phenomenon.

However, there is one further issue that arises from the variable effect of boundary strength across strong boundaries. The scaling of the F0 H pre-boundary peak across conditions BC2 to BC4 was found to be gradiently lower. According to the intonational approach, conditions BC2 to BC4 either belong to the same prosodic constituent (i.e. intermediate phrases) and are marked as such, or some conditions are intermediate phrases and others intonational. The finding of gradually lower scaling across BC2 to BC4 does not offer support for either view. If some conditions belong to intermediate phrases and others to intonational, no categorical effect was identified to support this claim. On the contrary, the alignment patterns were the same across conditions, and the scaling varied gradiently suggesting all three

conditions mirror different boundary strengths within the same domain. If, on the other hand, these variable differences stem from one type of prosodic domain, the question is whether they will be viewed as random variability apparent in the phonetic output, or whether they are viewed as anticipated boundary strength differences stemming from a linguistic variable that can predict this difference in boundary strength. In the present case the effects are due to manipulating syntax and length, they are apparent across speakers, and across phenomena. This suggests that this type of variability is best viewed as one that needs to be accounted for in some way.

It could also be suggested that the findings from BC0 in post-nasal stop voicing point to the existence of variability within the lowest domain. It was seen that within BC0 there were differences in how many times voicing occurred across items, and in how the items that did not reach agreement across raters were produced. This could suggest that there is variability also within the low levels of the structure, which should be investigated further.

Representing differences of boundary strength within the same domain type is not possible for labelled views. There is no available way for representing the proposal that there is a qualitative distinction between low and high domains, with variability allowed within each of the proposed domain types. The Strict Layering Hypothesis (Selkirk 1984) is an inherent part of all labelled views and, among other issues, it states that recursion is not allowed within the structure. This means that it is not possible to allow for one domain type to branch into the same domain type, and hence representing solely differences of strength. In past literature, it was seen that Ladd (1988, 2008) has offered an alternative to accommodate both differences of strength and type within the structure, by proposing the Compound Prosodic Domains. These are domains composed from restructuring the immediately higher domain. The present data fit this proposal well, showing a mixture of differences of strength and type across boundary conditions, and showing how differences of boundary strength do need to be represented in the prosodic structure.

One issue with Compound Prosodic Domains is the question of ‘how much’ recursivity is allowed. For example, it can be that boundary strength can range

indeterminately between levels of the same domain, or it can be that recursion can only take place once. The present thesis has not tested this question, but the data from pre-boundary lengthening and F0 suggest a three-way distinction between conditions BC2 to BC4, whereby the strength effect is similar on three of them. This supports the view that boundary strength can range in a variety of levels within labelled domains, but further research is needed¹.

One note regarding the application of post-nasal stop voicing; it could be argued that post-nasal stop voicing is not a sandhi process, but the result of a quantal effect (Stevens 1989). While it was found that the instances having and not having undergone voicing were statistically and perceptually different, it could be argued that the articulation from a voiceless to a voiced sound ranges in a continuum of values, as would be expected in the case of coarticulation, but it is perceived as a categorical phenomenon due to a quantal effect. That is, a very small articulatory change in the continuum of changes can give rise to a large acoustic and perceptual difference. For example, one such effect could be the difference between complete and non-complete obstruction of the nasal cavity by the velum. This would possibly be in support also of the instances upon which raters disagreed as to whether they had undergone voicing or not. Therefore, the articulatory realization of the phenomenon needs further investigation, and it should be coupled with an investigation of the articulatory behaviour of the pre-boundary nasal, since it is believed that its deletion might also take place categorically (Arvaniti & Joseph 2004).

Proposing that the phenomenon might be the result of masked coarticulation could be extended to suggesting that the actual articulatory difference across boundary conditions is not qualitatively distinct, as we suggested it is, but rather that the qualitative difference reported is a masked variable boundary strength effect. There are two points to be made countering this argument. First, qualitative differences were also identified in the alignment and scaling of F0 patterns, a finding that cannot be attributed to a quantal effect. Second, even if the effect of stop voicing is the result of a quantal effect, it is also possible that speakers are aware of the distinction between producing a voiceless or voiced stop, and make use of the phenomenon in

¹See also Kabak & Revithiadou (2007) for a phonological discussion on allowing recursivity of the Prosodic word; the present thesis mainly focused on higher level domains.

signalling prosodic domains (Stevens & Keyser 2008). The possibility of having a quantal effect in post-nasal stop voicing in Modern Greek is very interesting, and could be the matter of future research.

7.5 Length effects

One of the boundary conditions created for this thesis manipulated the length of the relevant utterances. To recapitulate the constructions, BC4 was a longer version of BC3 constructed by augmenting the number of syllables before and after the boundary. BC3 consisted of a short subordinate and main clause, with the boundary located between the two clauses. Hence, BC4 had the same syntactic construction but longer clauses. While the difference between BC3 and BC4 was one of length, the difference between BC2 and BC3 was one of syntax; BC2 was one clause and the boundary fell between a prepositional and a verb phrase, which was anticipated (intuitively and on the basis of Nespor and Vogel's analysis) to be a weaker boundary.

The expectation from the point of view of a labelled view would be that BC2 would be categorically distinct from BC3, while BC3 would not be different to BC4 in terms of their boundary strength. To illustrate this, the anticipated prosodic structure of the three conditions is shown:

BC2 (Prepositional phrase # Verb Phrase) $\rightarrow [\varphi \]\varphi \# [\varphi \]\varphi$

BC3 (Subordinate Clause # Main Clause) $\rightarrow [IP \]IP \# [IP \]IP$

BC4 (long Sub. Clause # long Main Clause) $\rightarrow [IP \]IP \# [IP \]IP$

The difference in boundary type across BC2 and BC3 is a qualitative difference of distinct domain type, while no difference is expected across BC3 and BC4. However, results from pre-boundary durations, vowel hiatus resolution and F0 scaling showed that this was not the case, and all three conditions ranged in a continuum of gradations. The difference across all conditions was similar, pointing to a difference of strength, while no qualitative effects were found. Two important conclusions can be drawn with respect to the effect of length on prosodic grouping. First, manipulating the length of the utterances was shown to have an effect on the strength, not the

type, of the relevant boundary. This suggests that length is a factor that needs to be taken into consideration when proposing theories of prosodic constituency, predicting prosodic domains. Moreover, the effect was shown to be similar to that of syntax in the immediately lower condition. That is, the magnitude of pre-boundary syllable lengthening was similar across BC2 and BC3, and across BC3 and BC4.

However, further research is needed in order to test the hypothesis that length effects can only give rise to differences of strength. The present thesis has only investigated one length manipulation, which altered the number of the relevant utterances quite abruptly. Further research could address more gradient changes of utterance length, testing differences of boundary strength and its interplay with syntactic constructions. The present results clearly indicate that length has not been adequately dealt with in previous research within labelled views.

7.6 Concluding Remarks

The present thesis has investigated boundary strength effects on a number of acoustic correlates to prosodic structure in an effort to understand whether the structure comprises of a set of labelled domains, solely differences of boundary strength, or a mixture of labelled domains with recursion. It has used instrumental data on phenomena previously analysed using impressionistic analyses, providing evidence in support and against labelled views. It has shown that labelled domains can be identified on the basis of qualitative boundary strength effects, and that also variable boundary strength differences need to be accounted for within the structure (by means of compound prosodic domains, indeterminate recursion, or any other similar suggestion). Additionally, it was shown that a theory predicting prosodic constituency needs to take into account a variety of factors; aside from the concurrent existence of qualitative and variable differences across levels, prosodic theories need to account for the effect of not only syntax, but also of length, on grouping phrases to prosodic constituents. The computation of how prosodic constituents are constructed is based on a variety of interacting factors, some of which were addressed in this thesis and were shown to significantly alter both the type and the strength of prosodic boundaries. This thesis has provided reliable data in support

of the concurrent existence of differences of domain *type* and *strength* within the prosodic structure.

APPENDIX A

Materials

BC0

Duration

1. το παζάρι
/to pa'zari/
'the bargain'
2. το πειστήριο
/to pi'stirio/
'the evidence'
3. το παγκάρι
/to pa'gari/
'the poor-box'
4. τα καλέσματα
/ta ka'lezmata/
'the callings'
5. τα παζάρια
/ta pa'zarja/
'the bargains'
6. τα ταξίδια
/ta ta'ksiðia/
'the travels'

7. το πηδάλιο
/to pi'dalio/
'the rudder'
8. τα κοιτάσματα
/ta ki'tazmata/
'deposits'

Vowel Hiatus

ao

1. τα ομόλογα
/ta o'mologa/
'bonds'
2. τα οπωρά
/ta o'pora/
'groceries'
3. τα ομόηχα
/ta o'moixa/
'homophonous'
4. τα ομόμετρα
/ta o'mometra/
'omometers'
5. τα οπάλια
/ta o'palia/
'opals'
6. τα οπτικά
/ta opti'ka/
'specs'

oa

1. το απόκρυφο
/to a'pokrifo/
'secret'

2. το απόκτημα
/to a'pɔktima/
'asset'
3. το απόγευμα
/to a'pɔɣevma/
'evening'
4. το απόβαρο
/to a'pɔvapo/
'dead-weight'
5. το αμόνι
/to a'mɔni/
'anvil'
6. το αμμόλουτρο
/to a'mɔlutro/
'sand-bath'

aa

1. τα απόβαρα
/ta a'pɔvapa/
'dead-weights'
2. τα απόνερα
/ta a'pɔnɛra/
'wake'
3. τα αμόνια
/ta a'mɔɲa/
'anvils'
4. τα απόκρυφα
/ta a'pɔkrifa/
'secrets'
5. τα αμμόλουτρα
/ta a'mɔlutra/
'sand-bath'

6. τα απόβλητα
/ta a'povlita/
'waste'

οο

1. το ομόλογο
/to o'mologo/
'the bond'
2. το οπάλιο
/to o'palio/
'opal'
3. το ομόρριζο
/to o'morizo/
'from same etymological root'
4. το ομόηχο
/to o'moixo/
'homophonous'
5. το ομόμετρο
/to o'mometro/
'omometer'
6. το ομώνυμο
/to o'monimo/
'the homonym'

Post-nasal stop voicing

1. όταν πακετάρεις
/'otan pace'taris/
'when (you) pack'
2. αν πακετάρεις
/an pace'taris/
'if (you) pack'

3. δεν πακετάρεις
/ðen pake'taris/
'do not pack'
4. μην πακετάρεις
/min pake'taris/
'do not pack'
5. τον πακετάρεις
/ton pake'taris/
'(you) pack him'
6. όταν ταξιδεύεις
/'otan taksi'ðevis/
'when (you) travel'
7. αν ταξιδεύεις
/an taksi'ðevis/
'if (you) travel'
8. δεν ταξιδεύεις
/ðen taksi'ðevis/
'(you) do not travel'
9. μην ταξιδεύεις
min taksi'ðevis
'do not travel'
10. την ταξιδεύεις
/tin taksi'ðevis/
'(you) travel her'
11. όταν καταγγείλεις
'otan kata'gilis
'when (you) denounce'
12. αν καταγγείλεις
/an kata'gilis/
'if (you) denounce'

13. δεν καταγγείλεις
ðen kata'gilis
'(you) do not denounce'
14. μην καταγγείλεις
/min kata'gilis/
'do not denounce'
15. τον καταγγείλεις
/ton kata'gilis/
'(you) denounce him'

BC1 to BC5

The same items were used throughout all boundary conditions. For an example of sentences constructed, see text in Chapter 2.

Duration

1. ασύδοτο παζάρι
[a'siðoto pa'zari]
'lawless bargain'
2. ασύδοτα παζάρια
[a'siðota pa'zarja]
'lawless bargains'
3. ανίκητο πειστήριο
[a'nicito pi'stirio]
'irrefutable proof'
4. αδιάκοπα καλέσματα
[a'ðjakopa ka'lezmata]
'endless shouting'
5. ψεύτικο παγγάρι
[pseftiko pa'gari]
'fake poor-box'
6. αφύσικα ταξίδια
[a'fisika ta'ksiðja]
'unnatural travels'

7. κοινότοπη κοπέλα
[ci'notopi ko'pela]
'common girl'
8. πολύπλοκη κατάσταση
[po'liploci ka'tastasi]
'complicated situation'

Vowel Hiatus

ao

1. αδύνατα ομόλογα
[a'ðinata o'moloɣa]
'weak bonds'
2. κατώτατα ομόλογα
[ka'totata o'moloɣa]
'lowest bonds'
3. ανάκατα ομόλογα
[a'nakata o'moloɣa]
'mixed bonds'
4. θεόρατα ομόλογα
[θe'orata o'mometra]
'huge omometers'
5. θεόρατα οπάλια
[θe'orata o'palia]
'huge opals'
6. αδύνατα οπτικά
[a'ðinata opti'ka]
'weak specs'

oa

1. υπέρτατο απόφθευγμα
[i'pertato a'pofθevɣma]
'supreme saying'

2. ανφτατο απόκτημα
[a'notato a'poktima]
'supreme asset'
3. ενδέκατο απόστολο
[en'dekato a'postolo]
'eleventh apostle'
4. αδύνατο απόγονο
[a'dinato a'royono]
'weak descendant'
5. αδύνατο αμόνι
[a'dinato a'moni]
'weak anvil'
6. πρόσφατο αμμόλουτρο
['prosfato a'molutro]
'recent sand-bath'

aa

1. ανώτατα αποκτήματα
[a'notata apo'ktimata] 'supreme assets'
2. υπέρτατα αποφθεύγματα
[i'pertata apo'fθevymata]
'supreme sayings'
3. αδύνατα αμόνια
[a'dinata a'monia]
'weak anvils'
4. κατώτατα αποφθεύγματα
[ka'totata apo'fθevymata]
'lowest sayings'
5. πρόσφατα αμμόλουτρα
['prosfata a'molutra]
'recent sand-bath'

6. ανώτατα αμορτισέρ
[a'notata amorti'ser]
'best shock-absorbers'

οο

1. αδύνατο ομόλογο
[a'dinato o'mologo]
'weak bond'
2. κατώτατο ομόλογο
[ka'totato o'mologo]
'lowest bond'
3. ανώτατο σπλίτη
[a'notato o'pliti]
'topmost private'
4. πρόσφατο οπωρώνα
[prosfato opo'rona]
'new orchard'
5. θεόρατο ομώμετρο
[θe'orato o'mometro]
'huge omometer'
6. θεόρατο σπάλιο
[θe'orato o'palio]
'huge opal'

Post-nasal Stop Voicing

1. αδάπανων ταμπέλων
[a'ðapanon ta'belon]
'free signs'
2. παράνομων πυλώνων
[pa'ranomon pi'lonon]
'illegal polls'
3. πτυσσόμενων πακέτων
[pti'somenon pa'ceton]
'foldable packs'
4. ακίνδυνων παράσιτων
[a'cinðinon pa'rasiton]
'harmless parasites'
5. απρόσμενων κοιτίδων
[a'prozmenon ci'tiðon]
'unexpected starting points'
6. χαρμόσυνων καμπάνων
[xar'mosinon ka'banon]
'happy-sounding bells'
7. αδύναμων τυφώνων
[a'ðinamon ti'fonon]
'weak typhoon'
8. ακίνδυνων καζίνο
[a'cinðinon ka'zino]
'harmless casinos'

APPENDIX B

F0

Figure B.1 shows a breakdown of the intonation contour before and after the boundary for each speaker, showcasing how BC1 was always differently produced to the other conditions.

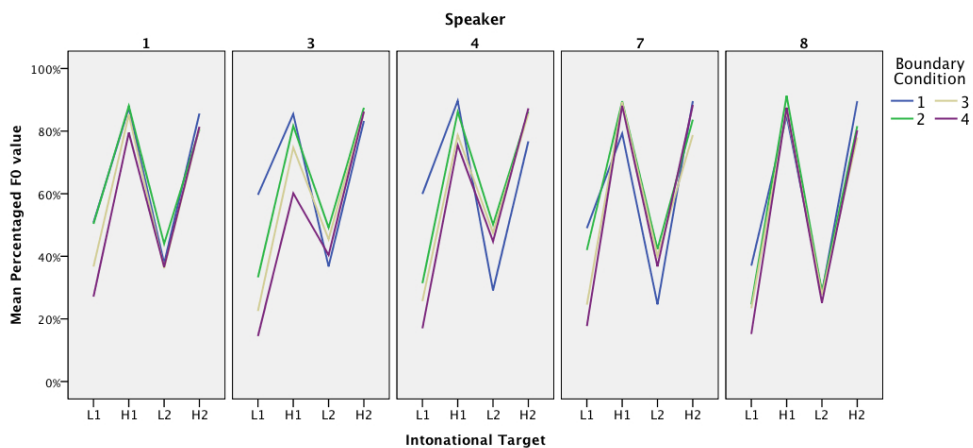


Figure B.1: Mean range values for each speaker separately for each measurement point. Duration is not represented in this graph.

APPENDIX C

Stop Voicing

		Voiceless	Voiced	Significant
Duration Closure	k g	45	58	**
	t d	44	70	**
	p b	63	65	-
Duration Release	k g	29	16	**
	t d	23	13	**
	p b	18	11	**
Amplitude Closure	k g	-19	-3	**
	t d	-14	-2	**
	p b	-18	-3	**
Amplitude Release	k g	-12	-4	**
	t d	-10	-1	**
	p b	-9	-1	**

**p<.001

Table C.1: Table showing the mean values of duration and amplitude of closure and release for voiceless and voiced stops.

	95% CI for exp b			
	B(SE)	Lower	exp b	Upper
Included				
Constant	5.1* (3.9)	1.38	1.75	2.22
Closure Amplitude	0.56** (0.12)			
Closure Duration	80.95* (34.6)			

Note $R^2=.72$ (Cox & Snell), .94 (Nagelkerke)
 Model $\chi^2(4)=301.53$, $p<.001$
 * $p<.05$, ** $p<.001$

Table C.2: Results from Binary Logistic Regression analysis testing which predictor values from the duration and amplitude of the closure and release best predict the voicing contrast between voiceless and voiced stops.

	95% CI for exp b			
	B(SE)	Lower	exp b	Upper
Included				
Constant	1.66* (0.65)	1.15	1.23	1.31
Release Amplitude	0.21** (0.03)			
Closure Amplitude	0.31** (0.56)			
Closure Duration	0.24* (0.01)			

Note $R^2=.54$ (Cox & Snell), .76 (Nagelkerke)
 Model $\chi^2(4)=473.23$, $p<.001$
 * $p<.05$, ** $p<.001$

Table C.3: Results from Binary Logistic Regression analysis testing which predictor values from the duration and amplitude of the closure and release best predict the voicing contrast between 'nasal+stop' sequences perceived to entail either voiceless or voiced stops.

References

- Abney, S. (1992), Prosodic structure, performance structure and phrase structure, *in* ‘Human Language Technology Conference, Proceedings of the Workshop on Speech and Natural Language’, Vol. 425-428.
- Adank, P. M. (2003), Vowel normalization: a perceptual-acoustic study of Dutch vowels, PhD thesis, Katholieke Universiteit Nijmegen.
- 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’, *Ellinika Grammata*, Athens, pp. 383–390.
- Arvaniti, A. (2000), ‘The phonetics of stress in Greek’, *Journal of Greek Linguistics* **1**, 9–39.
- Arvaniti, A. (2001*a*), Comparing the phonetics of single and geminate consonants in Cypriot and Standard Greek, *in* ‘Proceedings of the Fourth International Conference on Greek Linguistics’, pp. 37–44.
- Arvaniti, A. (2001*b*), The intonation of wh-questions in Greek, *in* ‘Studies in Greek Linguistics’, Vol. 21, pp. 57–68.
- Arvaniti, A. (2003), Peak scaling in Greek and the role of declination, *in* ‘Proceedings of the 15th International Congress of Phonetic Sciences’, pp. 2269–2272.
- Arvaniti, A. (2007), ‘Greek phonetics: The state of the art’, *Journal of Greek Linguistics* **8**, 97–208.
- Arvaniti, A. & Baltazani, M. (2005), Intonational analysis and prosodic annotation of Greek spoken corpora, *in* S.-A. Jun, ed., ‘Prosodic Typology: The Phonology of Intonation and Phrasing’, Oxford University Press, pp. 84–117.

- Arvaniti, A. & Joseph, B. D. (1999), Variation in voiced stop prenasalization in Greek, *in* 'Ohio State Working Papers in Linguistics', Vol. 52, pp. 203–233.
- Arvaniti, A. & Joseph, B. D. (2004), 'Early Modern Greek /b d g/: evidence from Rembetika and Folk songs', *Journal of Modern Greek Studies* **22**, 73–94.
- Arvaniti, A. & Ladd, D. R. (1995), Tonal alignment and the representation of accentual targets, *in* 'Proceedings of the 13th International Congress of Phonetic Sciences', Vol. 4, pp. 220–223.
- Arvaniti, A., Ladd, D. R. & Mennen, I. (1998), 'Stability of tonal alignment: the case of Greek prenuclear accents', *Journal of Phonetics* **26**, 3–25.
- Auran, C. (2008), 'Momel/intshint'.
- URL:** <http://stl.recherche.univ-lille3.fr/sitespersonnels/auran/english/index.html>
- Baltazani, M. (2006), Focusing, prosodic phrasing, and hiatus resolution in Greek, *in* L. Goldstein, D. Walen & C. Best, eds, 'Proceedings of Laboratory Phonology VIII', Mouton de Gruyter, Berlin, pp. 473–494.
- Baltazani, M. (2007a), The effect of prosodic boundaries on syllable duration in Greek. Abstract from the Old World Conference in Phonology 4.
- URL:** <http://www.rhodes.aegean.gr/tms/OCP4/Speakers/named-mbaltazani.pdf>
- Baltazani, M. (2007b), On s-voicing in Greek, *in* 'Proceedings for the 7th International Conference on Greek Linguistics'.
- Baltazani, M. & Jun, S. (1999), Topic and focus intonation in Greek, *in* 'Proceedings of the 14th International Congress of Phonetic Sciences', Vol. 2, p. 1305.
- Beckman, M. E. (1986), *Stress and Non-Stress Accent*, Dordrecht: Foris.
- Beckman, M. E. (1996), 'The parsing of prosody', *Language and Cognitive Processes* **11(1/2)**, 17–67.
- Beckman, M. E. & Edwards, J. (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 University Press, pp. 152–178.
- Beckman, M. E. & Edwards, J. (1994), Articulatory evidence for differentiating stress categories, *in* P. A. Keating, ed., 'Phonological Structure and Phonetic Form; Papers in Laboratory Phonology III', Cambridge University Press.

- Beckman, M. & Elam, G. A. (1997), 'Guidelines for ToBI labelling'.
- Beckman, M. & Pierrehumbert, J. (1986), 'Intonational structure in Japanese and English', *Phonology Yearbook* **3**, 255–309.
- Boersma, P. (2001), 'Praat, a system for doing phonetics by computer', *Glott International Glott International*(5:9/10), 341–345.
- Botinis, A. (1989), *Stress and Prosodic Structure in Greek*, Lund University Press.
- Botinis, A., Fourakis, M. & Prinou, I. (1999), Prosodic effects on segmental durations in Greek, in 'Eurospeech '99', Vol. 6, pp. 2475–2478.
- Byrd, D. & Saltzman, E. (1998), 'Intragestural dynamics of multiple prosodic boundaries', *Journal of Phonetics* **26**, 173–199.
- Byrd, D. & Saltzman, E. (2003), 'The elastic phrase: Modeling the dynamics of boundary-adjacent lengthening', *Journal of Phonetics* **31**(2), 149–180.
- Calhoun, S. (2006), Information Structure and the Prosodic Structure of English: a Probabilistic Relationship, PhD thesis, University of Edinburgh.
- Cambier-Langeveld, T. (1997), The domain of final lengthening in the production of Dutch., in 'Linguistics in the Netherlands', Amsterdam:John Benjamins, pp. 13–24.
- Carletta, J. C. (1996), 'Assessing agreement on classification tasks: the kappa statistic', *Computational Linguistics* **22** (2), 249–254.
- Casali, F. (1997), 'Vowel elision in hiatus contexts: Which vowel goes?', *Language* **73**(3), 493–533.
- Chavarria, S., Yoon, T.-J. & Cole, J. and Hasegawa-Johnson, M. (2004), Acoustic differentiation of ip and IP boundary levels: Comparison of L- and L-L% in the Switchboard corpus, in 'Speech Prosody 2004'.
- Cho, T. (2004), 'Prosodically conditioned strengthening and vowel-to-vowel coarticulation in English', *Journal of Phonetics* **23**, 141–176.
- Cho, T. (2005), 'Prosodic strengthening and featural enhancement: Evidence from acoustic and articulatory realizations of /a,i/ in English', *Journal of the Acoustical Society of America* **117**(6), 3867–3878.
- Cho, T. & Jun, S.-A. (2000), Domain-initial strengthening as enhancement of laryngeal features: aerodynamic evidence from Korean, in 'UCLA Working papers in Phonetics', Vol. 99, pp. 57–69.

- Cho, T. & Keating, P. (2001), 'Articulatory and acoustic studies on domain initial strengthening in Korean.', *Journal of Phonetics* **29**, 155–190.
- Cho, T. & Ladefoged, P. (1999), 'Variation and universals in VOT: evidence from 18 languages', *Journal of Phonetics* **27**(2), 207–229.
- Cho, T. & McQueen, J. (2005), 'Prosodic influences on consonant production in Dutch: Effects of prosodic boundaries, phrasal accent and lexical stress', *Journal of Phonetics* **33**(2), 121–157.
- Cho, T., McQueen, J. M. & Cox, E. A. (2007), 'Prosodically driven phonetic detail in speech processing: The case of domain-initial strengthening in English', *Journal of Phonetics* **35**(2), 210–243.
- Choi, H. (2003), Prosody-induced acoustic variation in English stop consonants, *in* 'Proceedings of the 15th International Congress of Phonetic Sciences'.
- Chomsky, N. & Halle, M. (1968), *The Sound Pattern of English*, New York: Harper and Row.
- Cole, J., Kim, H., Choi, H. & Hasegawa-Johnson, M. (2007), 'Prosodic effects on acoustic cues to stop voicing and place of articulation: Evidence from Radio News speech', *Journal of Phonetics* **35**, 180–209.
- Condoravdi, C. (1990), Sandhi rules of Greek and prosodic theory, *in* S. Inkelas & D. Zec, eds, 'The Phonology-Syntax Connection.', University of Chicago Press, pp. 63–84.
- Connell, B. (2001), Downdrift, downstep and declination, *in* 'Workshop on Typology of African Prosodic Systems', Bielefeld University, Germany, <http://www.spectrum.uni-bielefeld.de/TAPS/Connell.pdf>.
- Connell, B. (2004), Tone, utterance length and F0 scaling, *in* 'Proceedings of the International Symposium on Tonal Aspects of Languages: With emphasis on Tone Languages', ISCA, pp. 41–44.
URL: http://www.isca-speech.org/archive/tal2004/tal4_041.pdf
- Cooper, W. E. & Paccia-Cooper, J. (1980), *Syntax and Speech*, Cambridge MA: Harvard University Press.
- Cooper, W. E. & Sorensen, J. M. (1981), *Fundamental Frequency in Sentence Production*, Springer Verlag.

- de Pijper, J. R. & Sanderman, A. (1994), 'On the perceptual strength of prosodic boundaries and its relation to suprasegmental cues', *Journal of the Acoustical Society of America* **96**(4), 2037–47.
- del Mar Vanrell Bosch, M. (2007), Effect of utterance length on F0 scaling, in W. J. Barry & J. Trouvain, eds, 'Proceedings of the 16th International Congress of Phonetic Sciences', pp. 1165–1168.
- Dilley, L., Shattuck-Hufnagel, S. & Ostendorf, M. (1996), 'Glottalization of word-initial vowels as a function of prosodic structure', *Journal of Phonetics* **24**, 423–444.
- Drachman, G. & Malikouti-Drachman, A. (1999), Greek word accent, in H. van der Hulst, ed., 'Word Prosodic Systems in the Languages of Europe', Mouton de Gruyter, pp. 897–436.
- Eftychiou, E. (2007), Stop-Vowel coarticulation in Cypriot Greek, in W. J. Barry & J. Trouvain, eds, 'Proceedings of the 16th International Congress of Phonetic Sciences', pp. 517–520.
- Eftychiou, E. (2008), Lenition processes in Cypriot Greek, PhD thesis, Department of Linguistics, University of Cambridge.
- Ellis, L. & Hardcastle, W. J. (2002), 'Categorical and gradient properties of assimilation in alveolar to velar sequences: evidence from EPG and EMA data', *Journal of Phonetics* **30**, 373–396.
- Fallon, P. (1994), 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: John Benjamins Publishing Co, pp. 217–224.
- Féry, C. & Truckenbrodt, H. (2005), 'Sisterhood and tonal scaling', *Studia Linguistica* **59**(2), 223–243.
- Fougeron, C. (1999), Prosodically conditioned articulatory variations: A review, in 'UCLA Working papers in Phonetics', Vol. 97, pp. 1–74.
- Fougeron, C. (2001), 'Articulatory properties of initial segments in several prosodic constituents in French', *Journal of Phonetics* **29**(2), 109–135.

- Fougeron, C. & Keating, P. A. (1997), 'Articulatory strengthening at edges of prosodic domains', *Journal of the Acoustical Society of America* **101**, 3728–3740.
- Fourakis, M. (1986*a*), 'An acoustic study of the effects of tempo and stress on segmental intervals in Modern Greek', *Phonetica* **43**, 172–188.
- Fourakis, M. (1986*b*), 'A timing model for word-initial CV syllables in Modern Greek', *Journal of the Acoustical Society of America* **79**, 1982–1986.
- Fourakis, M., Botinis, A. & Katsaiti, M. (1999), 'Acoustic characteristics of Greek vowels', *Phonetica* **56**, 28–43.
- Frota, S. (1998), *Prosody and Focus in European Portuguese*, PhD thesis, Universidade de Lisboa.
- Gee, G. P. & Grosjean, F. (1983), 'Performance structures: A psycholinguistic and linguistic appraisal', *Cognitive Psychology* **15**, 411–458.
- Gibbon, F. & Scobbie, J. M. (1997), 'Covert contrasts in children with phonological disorder', *The Australian Communication Quarterly* **Autumn**, 13–16.
- Glasberg, B. R. & Moore, B. C. J. (1990), 'Derivation of auditory filter shapes from notched noise data', *Hearing Research* **47**, 103–137.
- Grice, M., Ladd, D. R. & Arvaniti, A. (2000), 'On the place of phrase accents in intonational phonology', *Phonology* **17**, 143–185.
- Grosjean, F. & Collins, M. (1979), 'Breathing, pausing and reading', *Phonetica* **36**, 98–114.
- Gussenhoven, C. (1999), 'Discreteness and gradience in intonational contrasts', *Language and Speech* **42**(2-3), 283–305.
- Gussenhoven, C. & Rietveld, A. (1997), Empirical evidence for the contrast between h* and l* in Dutch rising contours, in A. Botinis, G. Kouroupetroglou & C. G., eds, 'Intonation: Theory, models and applications', Athens, Grenoble: ESCA.
- Halliday, M. A. K. (1967), *Intonation and Grammar in British English*, Mouton.
- Hayashi, W., Hsu, C. & Keating, P. (1999), Domain-initial strengthening in Taiwanese: A follow-up study, in 'UCLA Working papers in Phonetics', Vol. 97, pp. 152–156.
- Hayes, B. & Lahiri, A. (1991), 'Bengali intonational phonology', *Natural Language and Linguistic Theory* **9**, 47–96.

- Hayes, B. & Stivers, T. (2000), Postnasal voicing. downloaded September 2007.
URL: <http://www.linguistics.ucla.edu/people/hayes/phonet/ncphonet.pdf>
- Hirst, D. (2007), A PRAAT plugin for MOMEL and INTSINT with improved algorithms for modelling and coding information, *in* W. J. Barry & J. Trouvain, eds, 'Proceedings of the 16th International Congress of Phonetic Sciences', pp. 1233–1236.
- Hirst, D. & Espesser, R. (1993), Automatic modelling of fundamental frequency using a quadratic spline function, *in* 'Travaux de l'Institut de Phonétique d'Aix', Vol. 15, pp. 75–85.
- Hockey, B. A. & Fagyal, Z. (1999), Phonemic length and pre-boundary lengthening: an experimental investigation on the use of durational cues in Hungarian, *in* 'Proceedings of the 14th International Congress of Phonetic Sciences', pp. 313–316.
- Hsu, C.-S. & Jun, S.-A. (1998), Prosodic strengthening in Taiwanese: syntagmatic or paradigmatic?, *in* 'UCLA Working papers in Phonetics', Vol. 96, pp. 69–89.
- Ito, J. & Mester, A. (1986), 'The phonology of voicing in Japanese: theoretical consequences for morphological accessibility', *Linguistic Inquiry* **17**, 49–73.
- Jun, S.-A. (1998), 'The Accentual Phrase in the Korean prosodic hierarchy', *Phonology* **15**, 189–226.
- Kabak, B. & Revithiadou, A. (2007), 'An interface approach to prosodic recursivity', Presentation at the 15th Manchester Phonology Meeting.
- Kainada, E. (2007), Prosodic boundary effects on duration and vowel hiatus in Modern Greek, *in* W. J. Barry & J. Trouvain, eds, 'Proceedings of the 16th International Congress of Phonetic Sciences', pp. 1225–1228.
- Kaisse, E. M. (1985), *Connected Speech: The Interaction of Syntax and Phonology*, Academic Press Inc.
- Katsika, A. (2007), Duration and pitch anchoring as cues to word boundaries in Greek, *in* W. J. Barry & J. Trouvain, eds, 'Proceedings of the 16th International Congress of Phonetic Sciences', pp. 929–932.
- Keating, P. A. (2003), Phonetic encoding of prosodic structure, *in* 'Proceedings of the 6th International Seminar on Speech Production', pp. 119–124.

- Keating, P., Cho, T., Fougeron, C. & Hsu, C. (2003), Domain-initial strengthening in four languages, *in* 'Papers in Laboratory Phonology VI'.
- Klatt, D. H. (1975), 'Vowel lengthening is syntactically determined in connected discourse', *Journal of Phonetics* **3**, 129–140.
- Kong, E., Beckman, E. & Edwards, J. (2007), Fine-grained phonetics and acquisition of Greek voiced stops, *in* '16th ICPHS', pp. 865–868.
- Kuzla, C. (2009), Prosodic Structure in speech production and perception, PhD thesis, Max Plank Institute of Psycholinguistics.
- Ladd, D. R. (1983), 'Phonological features of intonational peaks', *Language* **59**(4), 721–759.
- Ladd, D. R. (1986), 'Intonational phrasing: the case for recursive prosodic structure', *Phonology Yearbook* **3**, 311–340.
- Ladd, D. R. (1988), 'Declination 'reset' and the hierarchical organisation of utterances', *Journal of the Acoustical Society of America* **84**, 530–544.
- Ladd, D. R. (1996), *Intonational Phonology*, Cambridge University Press.
- Ladd, D. R. (2008), *Intonational Phonology*, 2nd edn, Cambridge University Press.
- Ladd, D. R. & Campbell, N. (1991), Theories of prosodic structure: evidence from syllable duration, *in* 'Proceedings of the 12th International Congress of Phonetic Sciences', pp. 290–293.
- Ladd, D. R. & Morton, R. (1997), 'The perception of intonational emphasis: continuous or categorical?', *Journal of Phonetics* **25**, 313–342.
- Ladd, D. R. & Scobbie, J. M. (2003), External sandhi as gestural overlap? Counter-evidence from Sardinian, *in* J. Local, R. Ogden & R. Temple, eds, 'Papers in Laboratory Phonology VI', Cambridge: Cambridge University Press, chapter 9, pp. 162–189.
- Landis, J. R. & Koch, G. G. (1977), 'The measurement of observer agreement for categorical data', *Biometrics* **33**(1), 159–174.
- Lengeris, A. (2009), 'Perceptual assimilation and L2 learning: evidence from the perception of Southern British English vowels by native speakers of Greek and Japanese', *Phonetica* **66**, 169–187.
- Lisker, L. & Abramson, A. S. (1964), 'A cross-language study of voicing in initial stops: acoustical measurements', *Word* **20**, 384–422.

- Lisker, L. & Abramson, A. S. (1967), The voicing dimension: some experiments in comparative phonetics, *in* 'Proceedings of the 6th International Congress of Phonetic Sciences', Academia Prague, pp. 563–567.
- Mackridge, P. (1985), *The Modern Greek Language*, Oxford University Press.
- Malikouti-Drachman, A. (2001), 'Greek phonology: A contemporary perspective', *Journal of Greek Linguistics* **2**, 187–243.
- Malikouti-Drachman, A. & Drachman, G. (1992), Greek clitics and Lexical Phonology, *in* W. U. Dressler, H. C. Luschutzky, O. E. Pfeiffer & J. R. Rennison, eds, 'Phonologica 1988', Cambridge: Cambridge University Press, pp. 197–206.
- Mirambel, A. (1961), *Introduction au Grec Moderne*, G. P. Maisonneuve.
- Nakai, S., Kunnari, S., Turk, A., Suomi, K. & Ylitalo, R. (2008), 'Utterance-final lengthening and quantity in Northern Finnish', *Journal of Phonetics* **37**(1), 29–45.
- Nespor, M. (1987), Vowel degemination and fast speech rules, *in* 'Phonology 4', pp. 61–85.
- Nespor, M. & Vogel, I. (1986), *Prosodic Phonology*, Foris Publications.
- Newton, B. (1972), *The Generative Interpretation of Dialect: a Study of Modern Greek phonology*, Cambridge University Press.
- Nicolaidis, K. (2001), 'An electropalatographic study of Greek spontaneous speech', *Journal of the International Phonetic Association* **31**(1), 67–85.
- Nicolaidis, K. (2002), Durational variability in Consonant-Vowel-Consonant sequences in Greek: the influence of phonetic identity, context and speaker, *in* M. Makri-Tsilipakou, ed., 'Selected Papers on Theoretical and Applied Linguistics from the 14th International Symposium', pp. 280–294.
- Nolan, F. (1992), The descriptive role of segments: Evidence from assimilation, *in* G. J. Docherty & D. R. Ladd, eds, 'Papers in Laboratory Phonology II: Gesture, Segment, Prosody', Cambridge: Cambridge University Press, pp. 261–280.
- Nolan, F., Holst, T. & Kühnert, B. (1996), 'Modeling [s] to [ʃ] accommodation in English', *Journal of Phonetics* **24**, 113–137.

- Onaka, A. (2003), Domain-initial strengthening in Japanese: An acoustic and articulatory study, *in* 'Proceedings of the 15th International Congress of Phonetic Sciences', pp. 2091–2094.
- Pagoni-Tetlow, S. (1993), 'Modern Greek prenasalisation', *Studies in Greek Linguistics* **14**, 355–369.
- Pelekanou, T. & Arvaniti, A. (2001), Postlexical rules and gestural overlap in a Greek spoken corpus, *in* 'Proceedings of the 5th International Conference on Greek Linguistics'.
- Philippaki-Warbuton, I. (1992), *Eisagogi stin theoritiki glossologia - Introduction to theoretical linguistics*, Nefeli/Glossologia.
- Pierrehumbert, J. B. (1980), The Phonology and Phonetics of English Intonation, PhD thesis, Massachusetts Institute of Technology.
- Pierrehumbert, J. & Beckman, M. E. (1988), *Japanese Tone Structure*, Cambridge, Mass.: MIT Press.
- Pierrehumbert, J. & Steele, S. A. (1989), 'Categories of tonal alignment in English', *Phonetica* **46**, 181–196.
- Pierrehumbert, J. & Talkin, D. (1992), Lenition of /h/ and glottal stop, *in* G. J. Docherty & D. R. Ladd, eds, 'Papers in Laboratory Phonology II: Gesture, Segment, Prosody', Cambridge: Cambridge University Press, pp. 90–119.
- Port, R. (1979), 'The influence of tempo on stop closure duration as a cue for voicing and place', *Journal of Phonetics* **7**, 45–56.
- Price, M. J., Ostendorf, M., Shattuck-Hufnagel, S. & Fong, C. (1991), 'The use of prosody in syntactic disambiguation', *Journal of the Acoustical Society of America* **90**(6), 2956–1970.
- Prieto, P., d'Imperio, M. & Gili Fivela, B. (2005), 'Pitch accent alignment in Romance: primary and secondary associations with metrical structure', *Language and Speech* **48**(359-396).
- Selkirk, E. (1978), On prosodic structure and its relation to syntactic structure, *in* T. Fretheim, ed., 'Nordic Prosody 2', TAPIR, Trondheim.
- Selkirk, E. (1981), On the nature of phonological representation, *in* T. Mayers, J. Laver & J. Anderson, eds, 'The Cognitive Representation of Speech', pp. 379–389.

- Selkirk, E. (1984), *Phonology and Syntax: the Relation Between Sound and Structure*, The MIT Press.
- Shattuck-Hufnagel, S. & Turk, A. E. (1996), 'A prosody tutorial for investigators of auditory sentence processing', *Journal of Psycholinguistics Research* **25**(2), 193–246.
- Siegel, S. & Castellan, N. J. (1988), *Non parametric statistics for the behavioral sciences*, London: McGraw-Hill.
- Silverman, K., Beckman, M., Pitrelli, J., Ostendorf, M., Wightman, C., Price, P., Pierrehumbert, J. & Hirschberg, J. (1992), ToBI: A standard for labeling English prosody, in 'Proceedings of the 1992 International Conference on Spoken Language Processing', Vol. 2, pp. 867–870.
- Stevens, K. (1989), 'On the quantal nature of speech', *Journal of Phonetics* **17**(1), 3–45.
- Stevens, K. (1998), *Acoustic Phonetics*, The MIT Press.
- Stevens, K. & Keyser, S. J. (2008), 'Quantal theory, enhancement and overlap', *Journal of Phonetics* **in press**.
- Tabain, M. (2003), 'Effects of prosodic boundary on /aC/ sequences: Acoustic results', *Journal of the Acoustical Society of America* **113**(1), 516–531.
- Trautmüller, H. (1990), 'Analytic expressions for the tonotopic frequency scale', *Journal of the Acoustical Society of America* **88**, 97–100.
- Tserdanelis, G. (2004), The role of segmental sandhi in the parsing of speech: evidence from Greek, PhD thesis, Ohio State University.
- Turk, A. & Dimitrova, S. (2007), English phrasal stress targets multiple, optional lengthening sites, in W. J. Barry & J. Trouvain, eds, 'Proceedings of the 16th International Congress of Phonetic Sciences', pp. 1177–1180.
- Turk, A. E. & White, L. (1999), 'Structural influences on accentual lengthening in English', *Journal of Phonetics* **27**, 171–206.
- Turk, A. & Shattuck-Hufnagel, S. (2000), 'Word-boundary-related duration patterns in English', *Journal of Phonetics* **28**, 397–440.
- Turk, A. & Shattuck-Hufnagel, S. (2007), 'Multiple targets of phrase-final lengthening in American English words', *Journal of Phonetics* **35**, 445–372.

- Wagner, M. & Gibson, T. (2006), 'Prosodic boundary strength scaling', Abstract from the 10th Laboratory Phonology Conference.
URL: <http://aune.lpl.univ-aix.fr/labphon10/abstracts/205.pdf>
- Watson, D. & Gibson, E. (2004), 'The relationship between intonational phrasing and syntactic structure in language production', *Language and Cognitive Processes* **19**(6), 713–755.
- Wightman, C. W., Shattuck-Hufnagel, S., Ostendorf, M. & Price, M. J. (1992), 'Segmental durations in the vicinity of prosodic phrase boundaries', *Journal of the Acoustical Society of America* **91**, 1707–1717.
- Yoon, T.-J., Kim, H. & Chavarria, S. (2004), Local acoustic cues distinguish two levels of prosodic phrasing: speech corpus evidence, *in* 'Poster presented at the 9th Laboratory Phonology'.
- Zsiga, E. C. (1997), 'Features, gestures and Igbo vowel assimilation: an approach to the phonology/phonetics mapping', *Language* **73**(2), 227–274.