Stress and Prosodic Structure in Greek

A Phonological, Acoustic, Physiological and perceptual Study

Antonis Botinis

TRAVAUX DE L'INSTITUT DE LINGUISTIQUE DE LUND

Lund
University
Press

TRAVAUX DE L'INSTITUT DE LINGUISTIQUE DE LUND 22

Stress and Prosodic Structure in Greek

A Phonological, Acoustic, Physiological and perceptual Study

Antonis Botinis

Lund University Press

Lund University Press
Box 141
S-221 00 Lund
Sweden

Art nr 20137
ISSN 0347-2558
ISBN 91-7966-080-0
Chartwell-Bratt ISBN 0-86238-223-8

Printed in Sweden
Bloms Tryckeri AB
Lund 1989

Electronic reprint
Athens 2011

Preface

This is an electronic edition of my 1989 doctoral thesis in phonetics from the University of Lund, Sweden. The thesis was first published by Lund University Press and has been out of print for several years now. In the present edition, special attention has been paid so that both form and content are as close as possible to the original publication. However, minor editorial amendments were carried out as well as some renumbering and organization of figures and tables, resulting in some deviations from the first publication.

When I started working on this edition I thought it would be an easy task. As the diskettes from the first publication were distorted, the whole book was scanned and edited thoroughly. It all ended up requiring much labour and much more time than I would have thought. Recalling the first writing with the very first Apple Macintosh in mid 1980s, it is admirable how far authorship technology, among other functions facilitated by computers, has advanced during the elapsed time. One need only mention that, in those days, no phonetic fonts were available. I remember designing basic phonetic symbols in Mac Paint for the phonetic transcriptions of the speech material. Phonetic research was also a laborious enterprise with analog equipment and thus no possibility of listening to the signal or of different displays.

At the time of my studies and research in Sweden, I was not aware that anybody else was working in Greek phonetics. In a short visit at MIT, however, Kenneth Stevens informed me that another Greek colleague, Marios Fourakis, was active in the USA. Marios and I exchanged several letters and, when I moved to Athens in 1991, we started collaboration resulting in joint publications and conference participations. The study of Greek phonetics has made significant progress since those pioneering times and is steadily joining the international community. This is expected to increase our knowledge of fundamental aspects of Greek and promote language studies and language applications in general.

Athens, December 2011
Antonis Botinis

Acknowledgements

This study has been carried out at the Phonetics Department, Lund University.

Upon my arrival in Lund, I had the opportunity to meet and study under Bertil Malmberg and Kerstin Hadding, who were at that time heads of the Linguistics and Phonetics Departments respectively.

Anders Löfqvist and Sidney Wood have been my teachers and have introduced me to the basic concepts of Phonetics.

In the Friday research seminar at the Phonetics Department, I have had a good chance to clarify ideas and exchange experience with colleagues and numerous visiting researchers.

The engineers, Gustav Jönsson, Bengt Mandersson, Mats Eeg-Olofsson, Mats Andersson and Lars Eriksson kindly provided me with their technical assistance during the different phases of my work.

Jan-Olof Svantensson gave me useful statistical advice, Jean Jacques Berthout helped me with the manuscript and David House revised my English skilfully and made content comments.

Greek students in Lund and Athens volunteered for the experiments and tried hard to do their best.

Our debt for considerable scientific achievements in this Department goes to Eva Gårding, who created a whole school of prosodists. She earnestly encouraged me and shared with me her keen interest in Greek and contrastive prosody.

Gösta Bruce introduced me to prosody and guided me to the final stage of my research. His authoritative expertise both in theory and practice and his deep insight into the Greek paradigm has proved priceless to me.

My mother Maria, and my father Theodore did whatever a parent can possibly do in an everlasting process.

My sincere thanks
Antonis Botinis

Contents

0 Introduction

0.1 Background

Prosody has historically taken up a considerable part of the scientific study of language. The word itself goes back to early Greek philosophers who used the term "προσωδία" to refer primarily to tonal features which, although not used in orthography, formed a central part of the language' system. From the time of Plato and Aristotle, a high tone "ὀξύς" (acute, sharp) and a low tone "βαρύς" (grave, heavy) are reported as well as the combination of the high and low (falling) tone within a syllable which was referred to as "περισπώμενος" (circumflex, bent round), "ὀξύβαρυς" (acute-grave) or "δύτονος" (bitone). These tonal distinctions along with "πνεῦμα" (aspiration) were later (around 200 B.C) introduced by Greek grammarians as orthographic symbols, mainly for practical reasons since Greek was the means of communication among peoples with different language backgrounds. Both "προσωδία" and "τόνος" (tension) reflect the understanding the Greeks had about the prosodic structure of the language which was one of pitch rather than stress.

By the end of the 4th century A.D. major phonological distinctions had already ceased in Greek resulting in a change in the prosodic system. The tonal distinctions faded away in a long lasting process and the tonal accents turned into dynamic accents or stress. This change meant that tonal phenomena were neglected in the study of prosody, and stress was associated with rhythm and the metrical structure of the language. On the other hand, the prosodic markings, totally redundant except for stress distinction, were established and used in standard Greek orthography until the past decade when the monotonic system was officially adopted, i.e. an acute mark on the vowel of the stressed syllable.

In the present century, the study of spoken languages has attributed to prosody considerable importance and linguists have tried to incorporate it into linguistic theory. The European tradition distinguishes the phonemic elements from the prosodic elements (Martinet 1965). Stress, accent, tone, quantity, rhythm and intonation are referred to as prosodic features whereas in the American tradition the term suprasegmentals in opposition to segmentals has been in large use (Hockett 1942, Trager and Smith 1951, Lehiste 1970). Suprasegmentals are either defined as a set of features mentioned above or as features whose domain extends over more than one segment and are classified through a syntagmatic comparison whereas segmental features are classified through a paradigmatic comparison.

With rapid advances in technology, it has been possible for impressionistic prosody to be investigated experimentally. Thus, as early as in the 1950's works by Fry (1955, 1958) and Bolinger (1958) could establish the acoustic correlates of stress and test them perceptually. In the 1960's, experimental techniques made it possible to establish the main physiological correlates of stress through both direct and indirect methods (Ladefoged 1967, Lieberman 1967). In the 1970's extensive acoustic studies on a variety of languages, mainly European ones, were being carried out and the accumulated knowledge resulted in attempts to formalize rules and construct early models on prosody (Bruce 1977, Bruce and Gårding 1978, Thorsen 1978, Pierrehumbert 1980). In the 1980's, with the development of computers and the emergence of speech technology the importance of prosody for speech synthesis and speech recognition systems has been widely acknowledged (Lea 1980) as well as its complicated character which demands further understanding.

Prosody may have a linguistic function for different levels and languages. At the lexical level, it may have a distinctive function for tone languages (e.g. Chinese, Thai), accent languages (e.g. Swedish, Serbocroatian) and stress languages (e.g. Greek, Italian). At the phrase level, prosody may have a syntactic function and divide the utterance into parts which belong together (e.g. French). At the sentence level, prosody may have a semantic function and highlight the most important part of the utterance (e.g. the majority of the European languages). At the discourse level, prosody may divide a sample of utterances into topics and subtopics, attribute to them the appropriate semantic salience according to the demands of the discourse and organize the utterances into turn interplay parts (probably a universal phenomenon). Moreover, prosody may have a paralinguistic function and convey the speaker's individual characteristics, sex, age, sociocultural background as well as his emotions and attitudes.

The Phonetics Department in Lund has a long and unbroken tradition on the study of prosody. The tradition goes back to Bertil Malmberg who introduced Phonetics as a science in Lund. Bertil Malmberg (1955) investigated the Swedish word accents and tested their acoustic correlates perceptually.

Kerstin Hadding-Koch (1961) investigated the intonation structure of the Southern Swedish dialect and made essential contributions to the study of the perception of intonation in collaboration with Michael Studdert-Kennedy (1964, 1973).

Eva Gårding (1967) investigated juncture phenomena in Swedish and in collaboration with Gösta Bruce (Bruce and Gårding 1978) developed what was to be called the Lund Model of Prosody and applied it contrastively in a number of languages with different prosodic systems (Gårding 1987).

Robert Bannert (1976) investigated temporal phenomena in German, worked for nearly a decade with contrastive studies in prosody and established the basis for a model of German prosody (1985).

Gösta Bruce (1977) investigated the Swedish word accents in sentence perspective and made essential contributions to the Lund model of prosody (1982) as well as to the study of dialectal prosodic variations (1983) and the rhythmic structure of Swedish (1984).

Recently, Paul Touati (1987) made a contrastive study on the prosodic structure of Swedish and French within the framework of the Lund model of prosody.

As for now, the tradition goes on with studies on prosodic parsing of Swedish (Gösta Bruce, David House and Lars Eriksson) in collaboration with Stockholm (Francisco Lacerda and Björn Lindblom) as well as on contrastive and interactive prosody (KIPROS) and the discourse structure of Swedish, French and Greek (Gösta Bruce, Ursula Willstedt, Paul Touati and Antonis Botinis).

0.2 The Object of Study

This study is about the prosodic structure of Greek. Greek (Ελληνική) is spoken by some 12 million people, 10 million in Greece and the rest in different parts of the world, mainly in the United States, Canada and Australia. The language investigated is standard Greek with Athenian Greek as the norm of the current standard. Athens is the capital and biggest city of Greece, about 4 million inhabitants, as well as the main cultural, political and economic center of the country. Athenian Greek is the language that has been spoken for centuries and includes lexical elements and expressions from dialects all over Greece, a result of the population growth of Athens during the last decades.

Greek is a branch of the west group of languages of the Indo-European family. Greek has a long history ever since the second millennium B.C. and has been documented up to the present day. The Greek language today, i.e. Modern Greek (Νέα Ελληνική), is the present stage of development of the language from Ancient Greek (Αρχαία Ελληνική), i.e. from the earlier recorded times up to the 3rd century B.C. From Ancient Greek it was Attic that absorbed dialectal varieties and gave rise to Hellenistic Koine (Common Greek), the international language of the Hellenistic Era, i.e. from the 3rd century B.C. to the 4th century A.D., from which Modern Greek emerged through Byzantine (up to the 11th century) and Medieval Greek (up to the 17th century).

From the Hellenistic Era on, probably even earlier, the Greek language shows two styles, a formal style resembling to a certain degree the prestigious Attic, and a colloquial style which was the natural development of the language. The polarization of the language as a means of communication mainly into a written (formal) and a spoken (colloquial) style up to the last decade has been known as “Katharevousa” (puristic) and “Dimotikí” (demotic).

During the last decade (1976), the demotic style obtained official recognition and is the norm of the standard language. This standard language as found in everyday use in Athens has absorbed most of the Katharevousa and/or archaic as well as provincial lexical elements and expressions into a unified means of expression. Today, the polarization of the language has given way to a common form with various styles to fulfil the multi-communicative purpose of the language.

The main purpose of this study is two-fold. First, to investigate how stress and prosody are related to morphology and syntax/semantics in Greek. Second, to describe the prosodic structure in Greek concentrating on stress, and examine the acoustic, physiological and perceptual correlates. This has been possible by the “isolation method”, i.e. by examining each prosodic category (word stress, enclitic stress, sentence stress) in environments where the influence of other prosodic phenomena is minimal as well as examining the prosodic parameters (duration, voice fundamental frequency, intensity) in different environments with different degrees of interplay.

Every prosodic category has a linguistic function which is shared by both the speaker and the listener. Since speech is produced by the human vocal tract and is transmitted through the air, there must be fairly constant physiological and acoustic correlates in the encoding process on the part of the speaker as well as acoustic and perceptual ones in the decoding process on the part of the listener. Modern theories of speech perception emphasize the link between speech production and speech perception or between speech acoustics and speech perception. Without arguing for the merits of the one or the other approach, our study is mainly based on acoustics which the listener must have a direct access to in order to interpret the intended message. The acoustic parameters of the prosodic categories were subjected to perceptual testing to validate their linguistic classification.

The prosodic categories examined in this study are the basis on which the rhythmic and intonation structure of the language are organized. Since this phonetic investigation is, to our knowledge, the first major experimental work on Greek prosody, we felt that a deeper understanding of the prosodic categories on four levels of analysis would form the basis for approaching other aspects of Greek prosody as well.

0.3 Outline

The study is composed of four parts. In the first part (1) the phonological system of Greek prosody at the lexical level, word level, phrase and sentence levels is described and rules for assigning lexical stress, word stress, enclitic and sentence stress to the corresponding level are presented.

The second part (2) is made up of three acoustic experiments. The purpose of experiment I was to investigate the contribution of the three acoustic parameters of duration, voice fundamental frequency (F0) and intensity to the production of word and sentence stress. Experiment Π was to investigate the contribution of the three parameters to enclitic stress and to compare the acoustic manifestation of an enclitic structure to the one of a proclitic structure. Experiment III was to investigate the acoustic manifestations of a paratactic and an enclitic structure and to examine their relation to prosody.

The third part (3) concerns a physiological investigation of the variations of subglottal pressure (Ps) associated with word and sentence stress. The purpose of this experiment was to find out how the acoustic parameters co-vary with subglottal pressure and if subglottal pressure affects one or more acoustic parameters, and to what degree.

The fourth part (4) is made up of six perceptual experiments. The purpose of experiment I was to find out which of the acoustic parameters contributes most to the perception of word stress after focus. Experiment II was to find out which of the acoustic parameters contributes most to word stress perception before focus, experiment III was designed to test the perceptual relevance of F0 changes in the frequency dimension, and experiment IV was to test F0 changes in the time dimension. Experiment V was to find out which of the acoustic parameters contributes most to the perception of enclitic stress, and experiment VI was designed to test the perceptual relevance of F0 for enclitic stress.

Finally (5), a summary and a conclusion of this study are presented.

1 Phonological Study

1.0 Introduction

In the phonological part of this study, the prosodic system of Greek at the lexical, word, phrase, and sentence level is described and rules describing the prosodic categories of lexical, word, enclitic, and sentence stress are introduced. The hierarchical structure of prosody is emphasized in accordance with the levels of representation involved in the production of the corresponding prosodic categories. The function and interdependence of morphology, syntax, and semantics in relation to prosody are examined and questions such as the Trisyllabic Constraint ~ Monotonic Principle, proclitic ~ enclitic structure, focus ~ presupposition in Greek are discussed.

1.1 Word Level Prosody

Word stress in Greek, according to our present knowledge, is provided by the lexicon for a considerable part of the language's vocabulary and has a distinctive function in words belonging to the same grammatical category, i.e. nouns [ˈnomos~noˈmos] (law~county), verbs [ˈperno~perˈno] (to take~to pass), and adverbs [ˈpote~poˈte] (when~never), as well as in words belonging to different grammatical categories, i.e. [ˈmilo~miˈlo] (apple~to speak), [ˈpoli~poˈli] (town~very), and [miˈtera~miteˈra] (mother~sharp).

Greek has a "limited freedom of stress", according to which stress appears within a "stressable zone" comprising the three last syllables of the word (Garde 1968). Every word that belongs to a major part of speech, i.e. noun, adjective, verb, adverb, has only one word stress (i.e. Monotonic Principle) which appears on one of the last three syllables (i.e. Trisyllabic Constraint), regardless of the number of syllables the word is composed of: [ameriˈci] (America), [eˈlaða] (Greece), [ˈvelɣio] (Belgium); on the other hand, function words like articles, pronouns and particles are usually unstressed, and may appear stressed only in a phrase or a sentence context.

Apart from the Monotonic Principle and the Trisyllabic Constraint[1], it seems that any general rules to predict the position of word stress in Greek would exclude a considerable body of the language's vocabulary. Earlier studies on this issue are more categorical about the unpredictability of word stress. Thus Tsitsopoulos (1973, p. 5) reports:

> The general feeling, at least from a pedagogical viewpoint, is that stress in MG (Modern Greek) is unpredictable and must be learned "as an essential part of the word" (Kahane and Ward 1945, quoted in Sotiropoulos 1972, p. 49).

Although the language tends to preserve word stress on the syllable of its basic position[2], inflectional words such as the noun and the verb may reassign word stress to another syllable when declined, (Nominative) [o ˈanθropos] ~ (Genitive) [tu anˈθropu] (the man), (Present) [ˈɣrafo] ~ (Past) [ˈeɣrapsa] (to write). We may refer to the basic position of stress provided by the lexicon as "lexical stress" and whenever morphology reassigns lexical stress through derivation and inflection it will be referred to as "word stress", i.e. lexical stress is lexically determined whereas word stress is morpholexically determined.

In contrast to Classical Greek which is believed to have been an accent language (Tsitsopoulos 1973, Steriade 1988), i.e. different pitch manifestations over the same prominent syllable used to convey linguistic (lexical) information, Modem Greek has turned into a stress language, utilizing only the stressed~unstressed distinction at the lexical level. Similar to our stress~accent classification of languages, this distinction has been referred to as dynamic or expiratory ~ musical or melodic (Lehiste 1970) as well as stress-accent~non-stress-accent languages (Beckman 1986). The

[1]"Demotic (spoken) Greek still preserves ton nómon ths trisyllabías, or rather the effect of the law of the three syllables of Classical Greek." (sotiropoulas 1972, p. 27).

[2]Word stress in declined forms is outside the scope of this study and will not be treated here; for a description of the Greek paradigm see Koutsoudas (1962), Warburton (1970a), Sotiropoulas (1972), and Tsitsopoulos (1973).

binary nature of word stress, and for prosodic categories in general, has been recognized for a number of languages by Vanderslice and Ladefoged (1972), Bruce (1977), Gårding (1977a), Schane (1979), Botinis (1982), Thorsen (1982), Rischel (1983), Bannert (1985), and Strangert (1985), among others.

The grammatical category “word” has been described as a linguistic unit on the domain of which both segmental and prosodic rules operate to give well-formed words (Chomsky and Halle 1968, Liberman and Prince 1977, Selkirk 1980). In Greek, prosodically well-formed words are subjected both to the Monotonic Principle, and the Trisyllabic Constraint (Warburton 1970, Tsitsopoulos 1973, Malikouti-Drachman 1976).

1.1.1 Non-compounds

The morphological system of the language provides evidence for the Monotonic Principle and the Trisyllabic Constraint. The open form classes contrast with each other both formally and distributionally. The noun and the verb, the most extensively inflecting word classes, are composed of a sequence of morphemes - the stem and the terminal. The stem may appear in its allomorphic variation whereas the terminal is highly inflectional denoting gender, number, and case for the noun as well as mood, tense, number and person for the verb.

The noun and the verb, when declined, tend to preserve word stress in its basic position, i.e., the singular nominative for the noun and the first person, present singular for the verb. The noun (Sing.) [to ˈθima] ~ (Pl.) [ta ˈθimata] (the victim) belongs to the same declension as the noun [to ˈmaθima ~ ta maˈθimata] (the lesson). Lexical stress on [ˈθima] is on the same syllable in both the singular and the plural; on the other hand, [to ˈmaθima], with its lexical stress on the antepenultimate, will have to assign its word stress one syllable to the right in the plural form, since the plural morpheme {ta} is an extra syllable breaking the Trisyllabic Constraint. The verb [paraˈðinome] (surrender) will have to move its lexical stress to the right as well, to form the first plural person [paraðiˈnomaste], in order not to violate the Trisyllabic Constraint.

We can proceed to the assignment of word stress on the words [maˈθimata] and [paraðiˈnomaste] with some assumptions about the words’ morphological representation and the application of the Word Stress Rule which takes into consideration both Monotonic Principle and Trisyllabic Constraint:

RULE 1 (WSR) S → [WS]/S_0_(S(S))#

[N[Nˈmaθima]Nta]N	[V[Vparaˈðino]Vmaste]V (WSR)
[Nmaˈθimata]N	[Vparaðiˈnomaste]V
[maˈθimata]	[paraðiˈnomaste]

(WSR=Word Stress Rule, S=Syllable, WS=Word Stress, #=word boundary)

1.1.2 Compounds

Compounds, no matter how complex, constitute a word prosodically being subjected both to the Monotonic Principle and the Trisyllabic Constraint, i.e. the accentual unit (Martinet 1960) in Greek (like in Russian and Italian) is the word, in contrast to the Germanic languages in which the accentual unit is the lexeme and the synthetic elements may keep the stress they carry as independent words. A compound may appear (1) with lexical stress on the same syllable as one of the synthetic words, [kaˈli ˈmera ~ kaliˈmera] (fine day ~ good morning), [kaˈli ˈtixi ~ kaˈlotiçi] (good luck ~ lucky) or (2), with lexical stress on a syllable which is unstressed in the single word, [ˈaɣrio aˈbeli ~ aɣriˈabelo] (wild vineyard), [ˈaɣria ˈɣata ~ aɣriˈoɣata] (wild cat).

The meaning of a compound may be predictable from its component elements [ˈaɣrio aˈbéli ~ aɣriˈabelo] (wild vineyard) but need not be, cf. [kaˈli ˈmera ~ kaliˈmera] (good day ~ good morning). Moreover, two words joined together may form two compounds different both in lexical stress distribution and meaning, [ˈçoni ˈðromos ~ çoˈnoðromos] (snow way ~ snow trail) and [ˈçoni ˈðromos ~ çonoˈðromos) (snow way ~ snow hiker).

In the [çoˈnoðromos ~ çonoˈðromos] minimal prosodic pair the stress of the synthetic words leaves no trace in the prosodic representation of the compounds. It rather seems that lexical stress is provided by the lexicon taking into consideration only the Trisyllabic Constraint and not the internal prosodic structure of the words composing the compound. Consequently, since the compound's stress is given by the lexicon, it is reasonable for the synthetic words to reach the lexicon unstressed, saving the grammar from any kind of ad hoc destressing rules.

1.2 Phrase Level Prosody

1.2.1 Enclitic Structure

When we come across phrase level prosody we see that the Monotonic Principle does not hold; moreover, the Word Stress Rule is not enough to describe phrase stress distribution.

(1)
to ˈθima (the victim)
ta ˈθimata (the victims)
to ˈθima mu (my victim)
ta ˈθimaˈta mu (the victims)

(2)
to ˈmaθima (the lesson)
ta maˈθimata (the lessons)
to ˈmaθiˈma mu (the lesson)
ta maˈθimaˈta mu (the lessons)

(3)
ˈðose (give)
ˈðose mu (give me)
ˈðose ˈmu to (give it me)

(4)
paˈraðose (deliver)
paˈraðoˈse mu (deliver to me)
paˈraðoˈse mu to (deliver it to me)

In (1), to [ˈθima] belongs to the same declension as to [ˈmaθima]. To form the plural, the formative {ta} is added at the end of the noun. Word stress on [ˈθimata] remains in the basic position, i.e. the formative {ta} does not affect the stress pattern of the noun. In the case of [ˈmaθima], when the plural formative {ta} is added, word stress is moved to the next syllable, [maˈθimata], according to the Word Stress Rule.

In the noun phrase [to ˈθima mu], word stress is in the basic position, i.e. the enclitic [mu] has no effect on the word's stress pattern either[3]. The case of [to ˈmaθiˈma mu] breaks the Monotonic Principle; moreover, the Word Stress Rule does not apply. We may consider [to ˈmaθiˈma mu] as a noun phrase where word stress is four syllables to the left of the phrase boundary, [ta ˈθimaˈta mu] appears with two stresses where neither the plural formative {ta} nor the enclitic mu is a justification for the second stress.

If we compare [to ˈmaθima] with [ta ˈθimata], we see the same stress pattern; however, if the formative {ta} is added to [to ˈmaθima], we have [ta

[3] A "host" refers to a word to which a clitic attaches; if the clitic follows its host it is an enclitic whereas if it precedes its host it is a proclitic (Zwicky 1977).

maˈθímata] whereas adding the enclitic [mu] to [ta ˈθimata] we get [ta ˈθimaˈta mu”. [ta maˈθimata] is a noun phrase consisting of an article+a word whereas [ta ˈθimaˈta mu] is a noun phrase consisting of an article+a word+a pronoun referring to the noun [ta ˈθimata], i.e. the internal structure of the phrase determines the one stress pattern or the other.

[ta maˈθimaˈta mu] is a noun phrase consisting of an article+noun, [ta maˈθimata], and the possessive pronoun [mu]. [ta maˈθimata] is a well-formed word after the Word Stress Rule has been applied. When we have the enclitic [mu] added to the word, the Word Stress Rule does not operate to move the word stress one syllable to the right; instead, a second stress appears. We may call the second stress derived from the internal structure of the phrase, “Enclitic Stress”[4].

In (3), [ˈðose] is the imperative of the verb [ˈðino], and in (4) [paˈraðose] is a compound composed of the verb [ˈðino] and the preposition [para]. When the enclitic [mu] is added to the above verbs, we have [ˈðose mu] with no word stress change whereas in the second verb we get [paˈraðoˈse mu] with an enclitic stress on the right of the word stress. Once more, the enclitic stress has nothing to do with the enclitic [mu] as such but rather with the phrase boundary. [ˈðose ˈmu to] bears two stresses; Word stress on the first syllable and enclitic stress on the enclitic [mu], which normally appears unstressed; in [paˈraðoˈse mu to], word stress is on the antepenultimate whereas enclitic stress is on the ultimate. Thus, the enclitic elements [mu] and [to] do not have a direct effect on the stress pattern of the phrase as such, but rather they function as post-word unstressed syllables, in the context of which the application of the enclitic stress takes place.

It seems that whereas the Word Stress Rule takes into consideration the number of unstressed syllables from the end of the word, the Enclitic Stress Rule starts counting from the word stress; if there are more than two unstressed syllables on the right of the word stress in a phrase, an enclitic stress appears on the second unstressed syllable no matter whether it is a lexical unit or an enclitic one. Thus, we may say that both words and phrases are subjected to the Trisyllabic Constraint; word stress may move to the right with the application of the Word Stress Rule at the word level whereas

[4]The term “enclitic stress” is used instead of “phrase stress” in the first edition.

enclitic stress may appear with the application of the Enclitic Stress Rule at the phrase level. We can now formulate the Enclitic Stress Rule which operates at the phrase level to produce the enclitic stress.

$$\text{RULE 2 (ESR) } S \rightarrow [ES] \,/\, S_0 \begin{bmatrix} S \\ WS \end{bmatrix} S\,[\_\_\_]\,S(S)\,\Big]_{NP,\,VP,\,AP}$$

(ESR=Enclitic Stress Rule, S=Syllable, ES=Enclitic Stress, WS=Word Stress

Warburton (1970a) proposes the "Antepenultimate Rule" to operate at the word level and the "Extended Antepenultimate Rule" to operate at the phrase level to assign word stress and enclitic stress, respectively, although she does not differentiate between the two prosodic categories. In a later study, Warburton (1970b) replaces the Extended Antepenultimate Rule by a cyclic application of the Antepenultimate Rule; she then adds a rule to dissimilate adjacent stresses to the right. Malikouti-Drachman (1976), although she does not differentiate between word stress and phrase stress either, takes care of the enclitic stress by a secondary auxiliary rule, similar to Warburton's Extended Antepenultimate Rule and our Enclitic Stress Rule. Compare now Warburton's cyclic derivation of [to ˈmaθiˈma mu] with Malikouti-Drachman's and our non-cyclic one.

Warburton (1970b)	Malikouti-Drachman	Botinis
[[to ˈmaθima]mu] AR	[[to ˈmaθima]mu] Aux. R.	[[to ˈmaθima]mu] ESR
[to ˈmaˈθima mu] DR	[to ˈmaθiˈma mu]	[to ˈmaθiˈma mu]
[to ˈmaθiˈma mu]		

(AR=Antepenultimate Rule, DR=Dissimilation Rule, Aux. R.=Auxiliary Rule, ESR=Enclitic Stress Rule)

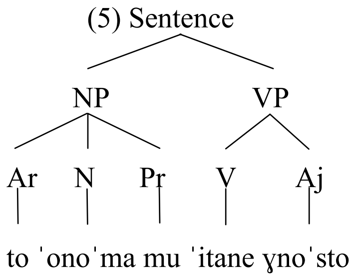
Warburton does not seem to take into account the fact that when the formative {ta} is added to [to ˈmaθima], [ta maˈθimata], we have derivational morphology whereas when the pronoun [mu] is added, [to ˈmaθiˈma mu], we have encliticization, two different processes. Both cases fulfil the natural bracketing hypothesis (Brame 1974) as well as the strict cyclicity principle (Kean 1974), since [to ˈmaθima] may appear as an independent word and the formatives {ta} and {mu} make crucial use of material in the outer cycle, hence evidence for cyclicity. [ta maˈθimata] and [to ˈmaθiˈma mu], in a cyclic application of the Antepenultimate Rule, reach the final cycle as [ta ˈmaˈθimata] and [to ˈmaˈθima mu]; to obtain the right prosodic output structure a destressing rule operates in the first case whereas a dissimilation rule operates in the second one, i.e. two different processes as the result of one rule, the Antepenultimate Rule. Malikouti-Drachman lets the Auxiliary Rule take care of this paradox which at the same time contributes to phonological simplicity since Warburton's dissimilation rule is eliminated from the grammar.

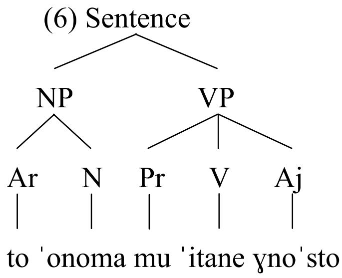
> When an unemphatic personal pronoun becomes enclitic to a preceding word (verb, noun, adverb) such that the three last syllables remain unstressed, the stress does not shift; rather a second stress arises, the assignment of which may lower the original stress. This fact will not be discussed, or taken into consideration here. (Malikouti-Drachman 1976, p. 107).

We think that Malikouti-Drachman is closer to linguistic reality. By letting the Word Stress Rule apply cyclically, we assume its products to be of the same nature. However, evidence has been provided (see part 2 & 4 of this study) that the word stress and the enclitic stress, apart from their phonological dissimilarity, have different acoustic and perceptual correlates. By adopting the non-cyclic application of prosodic categories we imply that stress rules run through a certain linguistic material only once. Although we still need information about word and phrase boundaries, it is the hierarchical structure among the same prosodic categories which is no longer necessary (Schane 1979).

1.2.2 Enclitic vs Proclitic Structure

In the minimal pair [to ˈonoˈma mu ˈitane ɣnoˈsto ~ to ˈonoma mu ˈitane ɣnoˈsto] (my name was familiar (to them) ~ the name was familiar (to me)) stress is distributed according to the whole phrase's internal structure. Although the two sentences have the same formatives, stress applies in a different way. Both sentences have the same word stress; moreover, the first sentence has an extra stress, an enclitic stress, as a result of the application of the Enclitic Stress Rule. If we examine the constituent structure, we see that the two sentences differ according to (5) and (6):





Considering the diagrams, we see that the clitic [mu] is dominated by a noun phrase in (5), but by a verb phrase in (6). On the other hand, [mu] is a possessive pronoun which modifies the preceding noun in (5) whereas in (6) [mu] is a personal pronoun, the indirect object of the verb, and does not have a direct syntactic relation to the preceding noun. Moreover, as an enclitic element, [mu] modifies the stress pattern of its host, whereas as a proclitic element it does not. Thus, clitics in Greek may be attached to a preceding host forming a syntactic and phonological enclitic structure or to a succeeding host to form a syntactic and phonological proclitic structure; clitics with a difference in syntactic and phonological attachment (Klavans 1985) have not been observed in Greek.

If we now topicalize the verb phrase in (5) and (6), we see that the enclitic [mu] is still an element of the noun phrase, whereas the proclitic [mu] is transformed at the beginning as a part of the verb phrase, [ˈitane ɣnoˈsto to ˈonoˈma mu ~ mu ˈitane ɣnoˈsto to ˈonoma]. Apart from the prosodic encliticization at the phrase level, there appear rules of segmental phonology applicable within the same domain (Selkirk 1980), i.e. voice assimilation, [o ˈanθroˈpo**z** mas | éɣrapse] ~ [o ˈanθropo**s** | mas ˈeɣrapse] (our man wrote ~ the man wrote to us).

1.3 Sentence Level Prosody

In everyday speech, a statement, from a communicative point of view, may be regarded as an answer to a possible question. The speaker wants to convey some information that the hearer is not supposed to know. The information already shared by the speaker and the hearer is the “presupposition” and the information the speaker wants to let the hearer have is the “focus” (Chomsky 1970, Jackendoff 1972).

Similar concepts to presupposition ~ focus distinction are theme ~ rheme (Danes 1960), [±mentioned] (Vanderslice and Ladefoged 1972) and [±highlighting] (Bolinger 1972); Rossi (1985) distinguishes between theme ~ rheme, an opposition au plan énontiatif (the pragmatic organisation of a sentence into given ~ new information), and topic ~ comment, an opposition au plan logico-semantique (the semantico-syntactic organisation of a sentence into something that one is talking about ~ what one says about it), a distinction of major significance when the two planes do not coincide.

In a Greek sentence like [to peˈði xtiˈpai ti ˈbala] (the child is kicking the ball), the word order is SVO; but even other syntactic structures associated with different contexts, more or less preferred, are possible like SOV, VOS, VSO, OSV, OVS. It seems as if the speaker has a free choice over which element of the sentence to put at the very beginning. However, the freedom of the speaker is drastically reduced to the element of the sentence that bears focus. Parts of the sentence that carry the most important information may be topicalized at the very beginning of the sentence in Greek.

On the other hand, the native speaker of Greek has another possibility to denote focus. The basic word order, i.e. SVO, may be kept and the main information may be conveyed by placing sentence stress, the prosodic category for focus, on one of the sentence’s elements. In this case sentence stress is semantically rather than syntactically conditioned (Danes 1960, Bolinger 1972, Jakendoff 1972, Bruce 1977).

1.3.1 Sentence Stress Distribution

Sentence stress is usually associated with the last element of the sentence in cases where we have a neutral sentence, i.e. a sentence without any contextual information (Chomsky and Halle 1968, Bruce 1977, Liberman and Prince 1977). However, in every day communication some parts of the sentence usually carry more important information than others and are consequently focused. In principle, any part of the sentence may be focused, i.e. sentence stress may appear at any position across the utterance as far as the syntactic and phonotactic conditions are not violated (Rossi 1985).

According to the Praguean Functionalist Sentence Perspective, a sentence is not regarded as an independent unit but as a part of a discourse (Danes

1960, 1967, 1974; Firbas 1974). The basic distinction is that a sentence from a contextual point of view is divided into a theme and a rheme, nearly corresponding to presupposition and focus, the rheme attracting sentence stress. A Greek sentence like [to miˈkro peˈði xtiˈpai ti ɟeˈnurja ˈbala] (the little boy is kicking the new ball) may appear in several contextual environments: (1) Q: [ti simˈveni] A: [to miˈkro peˈði xtiˈpai ti ɟeˈnurja ˈ**bala**] (What is happening? The little child is kicking the new **ball**). In this answer the whole sentence is the rheme and sentence stress will be on the rheme's last lexical element, **[ˈbala**] (bold letters indicate sentence stress). With a different context the word [ˈbala] may be the only lexical entity of the rheme and carry sentence stress: (2) Q: [pça keˈnurja xtiˈpai to miˈkro peˈði] A: [to miˈkro peˈði xtiˈpai ti ɟeˈnurja ˈ**bala**] (Which new is the little child kicking? The little child is kicking the new **ball**).

With the same word order but different contexts the sentence stress may be transferred to earlier sentence elements: (3) Q: [pça ˈbala xtiˈpai to miˈkro peˈði] A: [to miˈkro peˈði xtiˈpai ti **ɟenurja** ˈbala] (Which ball is the little child kicking? The little child is kicking the **new** ball). (4) Q: [ti ˈkani ti ɟeˈnurja ˈbala to miˈkro peˈði] A: [to miˈkro peˈði **xtiˈpai** ti ɟeˈnurja ˈbala] (What is the little child doing to the new ball? The little child **is kicking** the new ball). (5) Q: [pço miˈkro xtiˈpai ti ɟeˈnurja ˈbala] A: [to miˈkro **peˈði** xtiˈpai ti ɟeˈnurja ˈbala] (Which little is kicking the new ball? The little **child** is kicking the new ball). (6) Q: [pço peˈði xtiˈpai ti ɟeˈnurja ˈbala] A: [to **miˈkro** peˈði xtiˈpai ti ɟeˈnurja ˈbala (Which child is kicking the new ball? The **little** child is kicking the new ball).

However the theme, i.e. the old information, apart from its repetition devoid of sentence stress may be optionally pronominalized or omitted (Vanderslice and Ladefoged 1972). Sentence (6) may take the following answer forms: Q: [pço peˈði xtiˈpai ti ɟeˈnurja ˈbala] A1: [to **miˈkro** peˈðí xtiˈpai ti ɟeˈnurja ˈbala] ~ A2: [to **miˈkro** peðí ti xtiˈpai] ~ A3: [to **miˈkro** peðí] ~ A4: [to **miˈkro].**

In the last answer sentence, the rheme consists of only one word; in cases where the rheme is composed of several words the sentence stress appears on the last one (Danes 1960, Bruce 1977). (7) Q: [ti ˈkani to miˈkro peˈði] A: [to miˈkro peˈðí xtiˈpai ti ɟeˈnurja ˈ**bala**] (What is the little child doing? The little child is kicking the new **ball**). Although the whole verb phrase [xtiˈpai ti ɟeˈnurja ˈ**bala**] is the rheme, sentence stress appears on **[ˈbala**].

A break against contextuality may produce semantically anomalous sentences as in the following examples: (7) Q: [ti ˈkani to miˈkro peˈði] Al: *[to **miˈkro** peˈði xtiˈpai ti ɟeˈnurɟa ˈbala] ~ A2: *[to **miˈkro** peˈði ti xtiˈpai ~ A3: *[to **miˈkro** peˈði] ~ A4: *[to **miˈkro**]. Although the answers in (7) are completely well-formed in the context of (6), they become unacceptable in the context of (7), i.e. sentence stress is mainly contextually conditioned and its application is subjected to the sentence stress distribution rules.

The main elements of a sentence may be composed of several syllables each. Sentence stress will be mainly located on a constituent's particular syllable (see part 2), although its application may cause a reorganization of the acoustic (Rossi 1981) as well as the semantic structure (Bruce 1985) of the entire sentence. Given the focal information of a sentence the only thing we know is which element of the sentence, and not which particular syllable, will carry sentence stress. However, once the focal information of a sentence is given, sentence stress is bound to the sentence's morpholexical and syntactic structure (see part 2).

In simple sentences, the subject of this study, sentence stress will be on the rheme's last stressed syllable. However, the position of word stress within a word is determined by the lexical and morphological components. On the other hand, the position of enclitic stress is determined by the word stress and the syntactic component. Thus, sentence stress in Greek is conditioned by the utterance's contextual status at the semantic level and by the language's morpholexical and syntactic structure at the lower levels of representation.

1.4 Discussion

To recapitulate, in this study the following prosodic categories have been defined and described. First, lexical stress is the prosodic representation of a word at the lexical level. Second, word stress is the prosodic representation of a word at the morpholexical level. Third, enclitic stress is the prosodic representation of a phrase at the syntactic level and, finally, sentence stress is the prosodic representation of a sentence at the semanticontextual level.

Stress has traditionally been described as a one dimensional linguistic entity with a multivalued phonetic representation (Trager and Smith 1951, Chomsky and Halle 1968, Warburton 1970). In the present analysis, following the tradition of the prosodic school of Lund (Gårding 1973, 1977, Bruce 1977, Bannert 1982, Gårding et al. 1982), stress is described as a contribution of different prosodic categories with different linguistic functions which may merge into the same speech unit. The prosodic categories lexical stress, word stress, enclitic stress, and sentence stress have a classificatory function, i.e. either they exist on a certain speech unit or they do not (Thorsen 1982, Vanderslice and Ladefoged 1972 Rischel 1983). Through the different prosodic categories, the speaker of a particular language can make a sentence explicit given its morpholexical, syntactic, and semantic structure. Thus, "stress" may be considered as an abstract linguistic entity, a meeting point of the different levels of the language the domain of which varies from the word to the whole sentence.

Lexical stress is given by the lexicon and, apart from the Trisyllabic Constraint, does not take into account the higher levels of representation. In compounding, the synthetic words reach the lexicon as one unstressed lexical entity without taking into consideration any stress assignment in previous cycles as does cyclical phonology (Chomsky and Halle 1968, Brame 1974, Kean 1974) and cyclical metrical phonology (Kiparsky 1979). Selkirk (1980), in her description of English, eliminates the feature [±stress] below the word level, letting the prosodic structure itself be lexicalized; this is however superfluous for the description of Greek, which is lacking any other but the stressed ~ unstressed distinction.

English has been characterized as a stress-timed language (Lehiste 1977, Dauer 1983), i.e. stresses tend to appear at approximately equal time intervals; moreover, a stressed syllable which is metrically [S]trong may turn into [W]eak in certain environments under the influence of rhythm (Liberman and Prince 1977, Prince 1983, Hayes 1984). It seems that, in a stress-timed language, whenever two adjacent stresses are close together they may be dissimilated to comply with the language's rhythmic requirements. On the contrary, a stressed syllable in Greek may not undergo stress displacement since this may cause meaningful changes; moreover, there does not seem to be a pressure for such stress displacement since Greek may not be considered a strict stress-timed language (Dauer 1983).

Word stress may overlap with lexical stress across the Greek paradigm. However, through derivation and inflection the lexical stress may move to the right; given the lexical stress position, we assume the word stress distribution is provided by rules, a subject outside our study. Derivational and inflexional affixes are highly selective with respect to their hosts (Zwicky 1985) and may be thought of as word affixes operating at the word level which is subject both to the Monotonic Principle and the Trisyllabic Constraint. Thus, word stress is highly predictable given the lexical stress position and the word's morphological structure; the other levels of representation are not involved in its distribution and its relation to the other prosodic categories is the same as the one of lexical stress.

Enclitic stress, apart from the sentence's syntactic structure has to take into account the word stress involved in the enclitic structure as well, i.e. enclitic stress is bound to the phrase's lexical, morphological, and syntactic representation, the utterance's contextual structure being irrelevant for its application. By attributing enclitic stress to a prosodic category other than word stress where different components of the grammar with different functions are involved we no longer need the cycle (Chomsky and Halle 1968) or any kind of metrical grid (Liberman and Prince 1977) to account for its relative prominence over word stress within a given constituent; its prominence can be attributed to the fact that it belongs to a category with a higher position in the hierarchical structure of Greek prosody.

Klavans (1985) categorizes clitics as phrasal affixes. Given this definition we may say that in Greek, word affixes through derivation and inflexion may move word stress to the right whereas phrase affixes do not have this effect; they may instead produce the enclitic stress through encliticization, i.e. from morphology we have come to syntax. Apart from the different distribution and linguistic functions of word and enclitic stress we do have evidence from the acoustic and perceptual investigations that these two prosodic categories should be kept apart (see part 2 & 4).

Sentence stress, once the context has been determined, is distributed by rules (Gussenhoven 1983) on the last word of the focus domain. In elliptical sentences or in cases where the whole sentence is in focus, the neutral sentence included, the sentence stress will appear in final position. On the other hand, in cases where the presupposition is repeated, the focus units are usually topicalized at the beginning of the sentence in Greek, in contrast to the Slavonic languages where focus units often appear at the end of the sentence (Danes 1960).

A third possibility, the one examined in this study, is when the same word order of the response sentence is kept and by constructing appropriate contexts the presupposition ~ focus relation is changed, resulting in different sentence stress positions. However, apart from the contextual organization of the sentence, the sentence stress distribution rule has to take into account the lower levels of representation. Thus, sentence stress in Greek is highly predictable, given the contextual as well as the morpholexical and syntactic structure of a particular utterance. We are aware that different grammatical categories may effect sentence stress distribution (Home 1985) as well as pragmatics, an individual's expectations, beliefs or prejudices, a discussion we will not go into further.

To summarize, considering the relation of the prosodic categories to the levels of representation, we see that sentence stress, apart from the semantic division of the utterance into focus ~ presupposition, needs information about the syntactic as well as the morpholexical structure of the element it is to apply to (see part two). On the other hand, enclitic stress has to take into account only the syntactic and morpholexical structure of the phrase for its application. Finally, word stress is morpholexically conditioned whereas lexical stress is provided by the lexicon. In other words, it is only lexical stress which is relatively free within the last three syllables of a lexical word whereas the prosodic categories of enclitic stress and sentence stress are subjected to quite regular distribution rules.

2 Acoustic Study

2.0 Introduction

In the acoustic part of this study three experiments have been carried out. The purpose of the first experiment was to investigate the relative contribution of the acoustic parameters of duration, voice fundamental frequency (F0) and intensity to the production of word and sentence stress. The second experiment was to investigate the contribution of duration, F0 and intensity to enclitic stress and compare the acoustic manifestation of an enclitic structure with a word and an enclitic stress to the one of a proclitic structure with a word stress. The third experiment was to investigate the acoustic manifestation of two different syntactic structures, a paratactic structure with two word stresses to an enclitic structure with a word stress and an enclitic stress, and examine their relation to prosody.

2.1 Experimental Design

2.1.1 Subjects

The speakers of the speech material for the three acoustic experiments are five male students, in their early thirties at the time of the recording, brought up and educated in Athens; they speak what is considered to be standard Athenian, the core of the standard Greek language spoken in southern Greece, especially in Athens and Peloponnesos. None of the speakers has had any known history of speech, hearing, neurological, or respiratory disorders. All five speakers are monolingual with Greek as their mother language, speak about the same sociolect and are accustomed to their task as they have participated in similar phonetic experiments on several occasions.

2.1.2 Procedure

Proper meaningful Greek sentences were set up for the different purposes of the three acoustic experiments. The subjects were provided a set of paper cards with the test material for the experiments; each card contained only one sentence written in standard Greek orthography and prosodic markings.

For the neutral utterances, the test sentences were represented as simple statements without any contextual information; for utterances with different speech elements in focus, the test sentences were elicited as answers to a specific question having one sentence element in focus, the carrier of the most important information required by the question. The speakers were asked to read the cards eight times, each time in a different random order, as in everyday speech, with normal tempo and loudness. Each experiment took place on one occasion, lasted about fifteen minutes, and the subjects had no difficulties with their task apart from the second acoustic experiment where they made minimal focus production mistakes.

2.1.3 Acoustic Analysis

The test sentences were recorded on a Studer A62 tape recorder (ips 7.5) in a sound-treated room at Lund University Phonetics Laboratory. The frequency response of the tape recorder was flat within ±2 dB, from 30 to 14000 Hz, and the signal to noise ratio was 63 dB. The microphone was flat within the frequencies 35 to 17000 Hz.

The F0 contour was extracted by an F-J electronic pitch extracting device along with a duplex oscillogram and recorded on a Siemens oscillomink, with a paper speed of 100 mm/sec; the intensity contour was extracted by a Fonema analysis unit. F0 was calibrated in 10 Hz steps and intensity in 5 dB steps; both F0 and intensity scales were linear. In addition to the above mentioned analogue systems used in this study for the analysis of the main body of the data, a VAX 11/730 computer system with the ILS (api program) and some additional programs developed at Lund University Phonetics Laboratory were used as well. The accuracy of the measurements with the ILS program was within 1 Hz for F0, 1 dB for intensity, and 6,4 msec for duration; both F0 and intensity scales were linear.

The minima and maxima F0 values of the prosodic categories and structures under investigation were measured as well as the maxima intensity values; the duration of the segments, the syllables and, when necessary, the words or phrases were also measured. The intensity values are reported in decimals whereas the F0 and duration values in whole integers. Statistical calculations of the five speakers' mean, grand mean and standard deviations as well as two-tailed t-tests at 0.05 level of significance were carried out.

From the eight test sentence repetitions the first and the last ones were rejected; from the remaining six, the one which combined the least satisfactory production and/or acoustic display (in the investigator's judgement) was rejected too. The final five test sentences chosen provided the material for further acoustic analysis on which the results of this study are based.

In the present investigation, after a first examination of the acoustic manifestation of the prosodic categories in different contexts , the neutral as well as the entire focal utterances except for those in the first acoustic experiment were not considered further since they closely resembled the final focal utterances.

2.2 Acoustic Parameters of Word and Sentence Stress

2.2.1 Introduction

The main purpose of the first acoustic experiment was to investigate the contribution of duration, F0 and intensity to the production of word stress in prefocal, focal and postfocal position as well as the acoustic structure of sentence stress in Greek. Word stress is a prosodic category whose domain is the word and, apart from its culminative function in rhythmic structuring, has a distinctive function at lexical level and is morpholexically determined; sentence stress is a prosodic category whose domain is the sentence or parts of it, its function being to highlight the most important information of the sentence and is mainly semanticontextually determined (see part 1).

Thus, in our treatment of stress in Greek, the stressed syllable of a word is compared not only with its surrounding unstressed syllable(s) as being more prominent (syntagmatic plane) but also with its contrastive syllable(s) as being versus not being stressed (paradigmatic plane). Since word stress has a lexical function in Greek, a phonetic neutralization on the one plane may be compensated for by a distinction on the other plane, i.e. listeners may use any phonetic information available to them. Moreover, the stressed ~ unstressed distinction at the lexical level is compared with the focal ~ non-focal oppositions at the sentence level (contextual plane), i.e. the different prosodic organizations of the utterance according to external conditions which involves the requirements of the discourse.

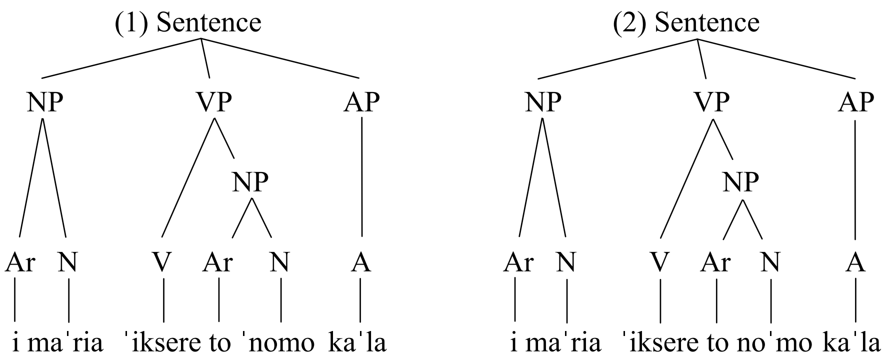
In this experiment three basic questions have been addressed: (1) What is the acoustic manifestation of word and sentence stress in Greek? (2) Which is the most constant acoustic parameter of word and sentence stress? (3) Are the acoustic parameters equally constant across the utterance or does their contribution to the production of word and sentence stress depend on the prosodic organization of the utterance? The acoustic parameters of duration, F0 and intensity are referred to as primary cues of stress (Lea 1977, Couper-Kuhlen 1986) whereas segmental cues like a glottal stop in English, vowel reduction in English and Russian, vowel deletion in some northern Greek dialects are referred to as secondary cues of stress.

However, none of the primary cues, nor the secondary ones, may be used in absolute terms but rather function in relative terms because of interfering factors: First, each vowel has its inherent duration, F0 and intensity (Lehiste 1970), i.e. low vowels may have longer duration and higher intensity but lower F0 than high vowels. These physiologically determined differences are not noticed in speech since the perceptual system compensates for this interference. Second, the segmental context may affect the acoustic values of a vowel (Di Cristo 1978), like voiceless consonants which cause a frequency jump on the following vowel whereas they reduce the duration of the preceding vowel. Third, F0 and intensity tend to decrease successively through the utterance (Pierrehumbert 1978) whereas syntactic finality tends to increase segmental duration (Paccia-Cooper and Cooper 1981).

From the primary cues, the importance of F0 over duration for stress distinction has been emphasized whereas the role of intensity has often been questioned. This has been suggested even for languages which would be classified as dynamic-stress languages like Danish (Thorsen 1982), English (Fry 1958, Bolinger 1958) and German (Bannert 1985) whereas for melodic-stress or pitch-accent languages like Swedish (Malmberg 1955, Bruce 1977) the importance of F0 has repetedly been validated.

2.2.2 Speech Material

Two meaningful Greek sentences were set up containing the prosodic minimal pair under investigation [ˈnomo~noˈmo] (law~county) in the carrier sentence [i maˈria ˈiksere to ___ kaˈla] (Maria knew the ___ well). The segmental structure of the minimal pair was composed exclusively of sonorants in order to have a continuous and easily detected F0 contour. Moreover, nasals which have a minimal influence on F0 and the mid, back vowel was chosen, to avoid intrinsic differences which vowels with different degrees of opening may have. Both sentences were declarative with the same syntactic structure, i.e. subject-verb-object-adverb, according to (1) and (2).



Each of the two test sentences was the carrier of five different contextual organizations listed in Table Π.I. First, they were pronounced neutrally, i.e. with no contextual information (la-2a). The next three productions were elicited as answers formulated in different ways to make the speaker choose one of the elements of the sentence as the focus and the carrier of the information required by the questioner. In productions lb~2b, the adverb [kaˈla] was in focus having [ˈnomo~noˈmo] in prefocal position, in lc~2c [ˈnomo~noˈmo] was in focus and, in ld~2d, [maˈria] was in focus, having [ˈnomo~noˈmo] in postfocal position. The last production was presented as new information, i.e. the whole sentence was in focus (le~2e).

Table II.I. The minimal pair [ˈnomo~noˈmo] in the carrier sentence [i maˈria ˈiksere to ___ kaˈla] with different contextual organizations (see text).

Contextual frame	Test sentences
1a. None	[i maˈria ˈiksere to ˈnomo kaˈla] (Maria knew the law well.)
1b. [pos ˈiksere i maˈria to ˈnomo] (How did Maria know the law?)	[i maˈria ˈiksere to ˈnomo **kaˈla**] (Maria knew the law **well**.)
1c. [ti ˈiksere i maˈria kaˈla] (What did Maria know well?)	[i maˈria ˈiksere to ˈ**nomo** kaˈla] (Maria knew the **law** well.)
1d. [pça ˈiksere to ˈnomo kaˈla] (Who knew the law well?)	[i **maˈia** ˈiksere to ˈnomo kaˈla] (**Maria** knew the law well.)
1e. [ti ˈneα] (What news?)	[**i maˈria ˈiksere to ˈnomo kaˈla**] (**Maria knew the law well**.)
2a. None	[i maˈria ˈiksere to noˈmo kaˈla] (Maria knew the county well.)
2b. [pos ˈiksere i maˈria to noˈmo] (How did Maria know the county?)	[i maˈria ˈiksere to noˈmo **kaˈla**] (Maria knew the county **well**.)
2c. [ti ˈiksere i maˈria kaˈla] (What did Maria know well?)	[i maˈria ˈiksere to **noˈmo** kaˈla] (Maria knew the **county** well.)
2d. [pça ˈiksere to noˈmo kaˈla] (Who knew the county well?)	[i **maˈia** ˈiksere to noˈmo kaˈla] (**Maria** knew the county well.)
2e. [ti ˈneα] (What news?)	[**i maˈria ˈiksere to noˈmo kaˈla**] (**Maria knew the county well**.)

2.2.3 Results

The set up of the material of the first acoustic experiment permits a syntagmatic and a paradigmatic as well as a contextual comparison of both segments and syllables composing the prosodic minimal pair. Thus, in the [ˈnomo~noˈmo] opposition, the stressed syllable (and its segments) of the first member of the pair can be compared both with the unstressed syllable of the same member (syntagmatic plane) and the unstressed syllable of the second member (paradigmatic plane) and, vice versa, the unstressed syllable of the second member of the pair can be compared with the stressed syllable of the same member as well as with the stressed syllable of the first member of the pair [ˈnomo~noˈmo]. Moreover, the stressed~unstressed opposition of both members of the pair can be compared with the prefocal, focal, and postfocal productions of the utterances (contextual plane).

As an example, one speaker's raw data of duration, F0, and intensity of the minimal pair [ˈnomo~noˈmo] at different sentence contexts is shown in Figures 2.1, 2.3, 2.5; the five speakers' mean values of the prosodic categories of word and sentence stress are shown in Tables 2.1-2.3 and their grand mean in Figures 2.2, 2.4, 2.6.

2.2.3.1 Duration

In Figure 2.1 one speaker's carrier sentences of [ˈnomo~noˈmo] in different contexts are shown. It seems that although there is a temporal reorganisation of the entire utterances as a result of sentence stress application, the duration pattern of the sentences is fairly constant: Stressed syllables are longer than unstressed syllables (lb~2b, ld-2d) and stressed syllables in focus are longer than stressed syllables out of focus (lc, 2c ~ lb, ld, 2b, 2d), which is a contribution of word and sentence stress respectively.

Prefocal distribution

In prefocal position (Table 2.1, 1b~2b), the five speakers' syntagmatic comparison of the stressed~unstressed syllable of the first member of the pair [ˈnomo~noˈmo] showed a significant difference of 15 ms (126~111 ms, $t(4)=2.8$, $p<0.025$); the nasals composing the compared syllables were of about the same duration (50~49 ms) whereas the stressed vowel was 14 ms longer than the unstressed one, (76~62 ms, $t(4)=8.3$, $p<0.0005$).

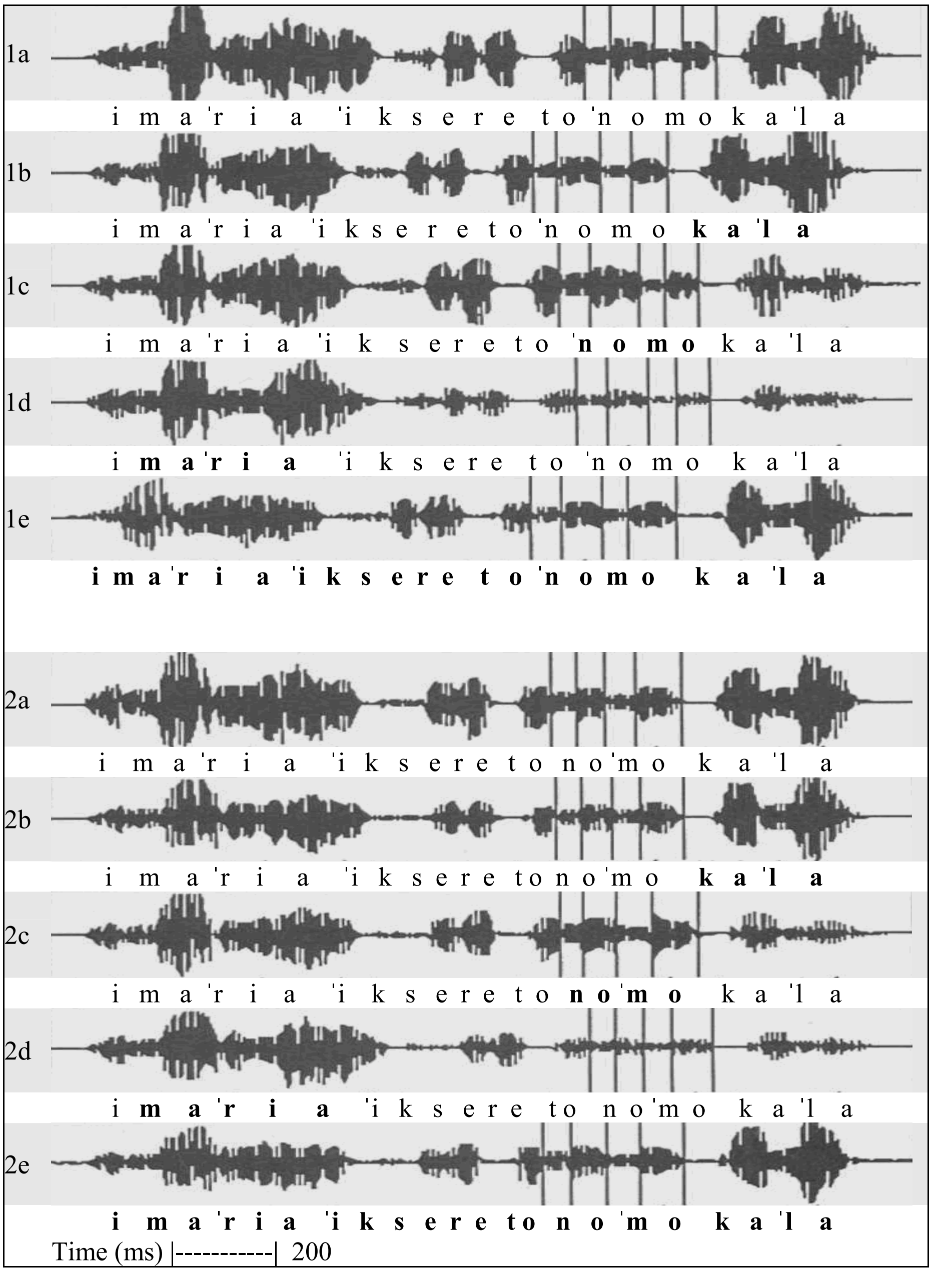


Figure 2.1. One speaker's raw data segmental durations of the minimal pair [ˈnomo~noˈmo] (1~2) in the carrier sentence [i maˈria ˈiksere to ___ kaˈla] with different focal organisations (a~e).

Table 2.1. Five speakers' mean segmental durations (ms), standard deviation (sd) and grand mean ($\bar{\bar{X}}$G) of the minimal pair [ˈnomo~noˈmo] in the carrier sentence [i maˈria ˈiksere to ___ kaˈla] in prefocal (lb~2b), focal (lc~2c), and postfocal position (ld~2d).

(lb~2b)	ˈn	o	m	o	n	o	ˈm	o
1 ms	58	78	62	70	54	62	62	80
sd	4.4	8.3	4.4	7.0	5.4	4.4	8.3	7.0
2	56	74	44	58	38	60	72	80
	5.4	5.4	5.4	4.4	8.3	7.0	8.3	7.0
3	36	74	52	56	40	58	54	78
	5.4	5.4	4.4	5.4	0.0	8.3	5.4	8.3
4	56	78	50	64	48	70	54	86
	5.4	4.4	7.0	5.4	4.4	7.0	5.4	5.4
5	46	76	40	62	34	56	56	76
	5.4	5.4	10.0	4.4	5.4	5.4	5.4	5.4
$\bar{\bar{X}}$G	50	76	49	62	42	61	59	80
sd	9.3	2.0	8.4	5.4	8.0	5.4	7.6	3.7
(lc~2c)	**ˈn**	**o**	**m**	**o**	**n**	**o**	**ˈm**	**o**
1 ms	56	96	62	68	58	64	62	98
sd	5.4	5.4	8.3	4.4	4.4	5.4	4.4	4.4
2	76	80	46	62	34	62	68	82
	11.4	7.0	8.9	8.3	5.4	4.4	4.4	8.3
3	58	88	52	58	40	64	58	90
	8.3	8.3	8.3	8.3	10.0	5.4	4.4	7.0
4	64	86	58	62	52	74	58	96
	5.4	5.4	4.4	4.4	4.4	5.4	4.4	5.4
5	52	88	38	66	34	68	58	90
	8.3	8.3	4.4	8.9	5.4	4.4	4.4	7.0
$\bar{\bar{X}}$G	61	87	51	63	43	66	60	92
sd	9.3	5.7	9.5	3.8	10.9	4.7	4.3	6.2
(ld~2d)	ˈn	o	m	o	n	o	ˈm	o
1 ms	56	84	58	68	50	60	66	82
sd	5.4	5.4	8.3	4.4	0.0	7.0	5.4	4.4
2	52	74	58	62	30	56	62	72
	4.4	5.4	8.3	4.4	0.0	5.4	4.4	4.4
3	48	78	54	52	42	66	66	82
	8.3	8.3	5.4	4.4	4.4	5.4	5.4	4.4
4	60	84	48	66	48	66	56	82
	7.0	5.4	8.3	5.4	4.4	5.4	5.4	8.3
5	42	74	52	60	34	66	46	84
	8.3	5.4	8.3	7.0	5.4	5.4	5.4	5.4
$\bar{\bar{X}}$G	51	78	54	61	40	62	59	80
sd	6.9	5.0	4.2	6.2	8.6	4.6	8.4	4.7

In the second member of the pair, the stressed syllable was 35 ms longer than the unstressed one, (139~103 ms, $t(4)=6.2$, $p<0.005$); both the nasal and the vowel of the stressed syllable were significantly longer than the ones of the unstressed syllable, 17 ms for the former (59~42 ms, $t(4)=3.2$, $p<0.025$) and 19 ms for the latter (80~61 ms, $t(4)=23.5$, $p<0.0005$).

The paradigmatic comparison of [ˈnomo~noˈmo] showed significant differences for both syllables and vowels but not for the nasals in this contrastive pair. The stressed syllable of the first member was 23 ms longer than the unstressed syllable of the second member (126~103 ms, $t(4)=4.4$, $p<0.01$), 8 ms for the nasal (50~42 ms, $t(4)=2.0$, $p>0.05$) and 15 ms for the vowel (76~61 ms, $t(4)=7.5$, $p<0.005$); the stressed syllable of the second member of the pair was 28 ms longer than the unstressed syllable of the first member (139~111 ms, $t(4)=4.3$, $p<0.01$), 10 ms for the nasal (59~49 ms, $t(4)=1.8$, $p>0.05$) and 18 ms for the vowel (80~62 ms, $t(4)=7.1$, $p<0.005$).

All speakers produced the majority of the stressed syllables and vowels of both members of the pair with longer duration whereas the nasals did not show this constancy. Although the nasals of the stressed syllables of the second member of the pair were longer than the nasals of the unstressed syllables as a rule, the nasals of the first member of the pair showed contradictory variations. Speakers 2, 4 and 5 produced the nasals of the stressed syllables with longer duration (Sp. 2, 56~44, Sp. 4, 56~50, Sp. 5, 46~40 ms) whereas speakers 1 and 3 produced them with shorter duration than in the unstressed syllables (Sp. 1, 58-62, Sp. 3, 36-52 ms).

Focal distribution

In focal position (Table 2.1, lc~2c), on the syntagmatic plane, the stressed syllables, especially the vowels, were significantly longer than the unstressed ones for both members of the [ˈnomo~noˈmo] pair. The stressed syllable of the first member of the pair was 34 ms longer than the unstressed one (148~114 ms, $t(4)=8.0$, $p<0.005$), 10 ms for the nasal (61~51 ms, $t(4)=1.6$, $p>0.05$) and 24 ms for the vowel (87~63 ms, $t(4)=11.4$, $p<0.0005$). The stressed syllable of the second member of the pair was 42 ms longer than the unstressed one (151~109 ms, $t(9.6)=p<0.0005$), 17 ms for the nasal (60~43 ms, $t(4)=3.0$, $p<0.025$) and 25 ms for the vowel (91~66 ms, $t(4)=9.9$, $p<0.0005$).

On the paradigmatic plane, the stressed syllable of the first member of the pair was 39 ms longer than the unstressed syllable of the second member of the pair (148~109 ms, t(4)=6.3, p<0.005), 18 ms for the nasal (61~43 ms, t(4)=2.4, p<0.05) and 21 ms for the vowel (87~66 ms, t(4)=6.3, p<0.005). The stressed syllable of the second member of the pair was 37 ms longer than the unstressed syllable of the first member (151~114 ms, t(4)=14.6, p<0.0005), 9 ms for the nasal (60~51 ms, t(4)=2.0, p>0.05) and 28 ms for the vowel (91~63 ms, t(4)=10.7, p<0.0005).

All speakers produced the focal stressed syllables and vowels of both members of the pair with longer duration than the unstressed syllables whereas the nasals did not show reliable constancy. The nasals of the stressed syllables of the second pair were longer than the nasals of the unstressed syllables whereas, for the first member of the pair, speaker 1 produced the nasal of the stressed syllable with shorter duration than the nasal of the unstressed syllable (56~62 ms).

Postfocal distribution

In postfocal position (Table 2.1, ld~2d), the duration pattern of both members of the pair [ˈnomo~noˈmo] was quite close to the prefocal one, i.e. the stressed syllables were significantly longer than the unstressed ones, the only difference between the prefocal ~ postfocal position worth mentioning was the postfocal, second nasal of the first member of the pair which, although in an unstressed syllable, was 3 ms longer than the first nasal of the same member (54-51 ms).

All speakers produced the stressed syllables and vowels of both members of the pair with longer duration than the unstressed syllables whereas the nasals did not show this constancy. Although the nasals of the stressed syllables of the second member of the pair were longer than the nasals of the unstressed syllables, the nasals of the first member of the pair showed the opposite variation for all speakers except speaker 4 who produced the nasals of the stressed syllables with longer duration than the nasals of the unstressed syllables (60~48 ms).

Contextual distribution

On the contextual plane (Figure 2.2), the prefocal and postfocal productions did not show any significant difference (at the 0.05 level) for both members of the pair [ˈnomo~noˈmo], their duration pattern being fairly the same. The pre(post)focal ~ focal comparison of the stressed syllables of the first member showed significant differences (126~148 ms, t(4)=-5.7, p<0.005),

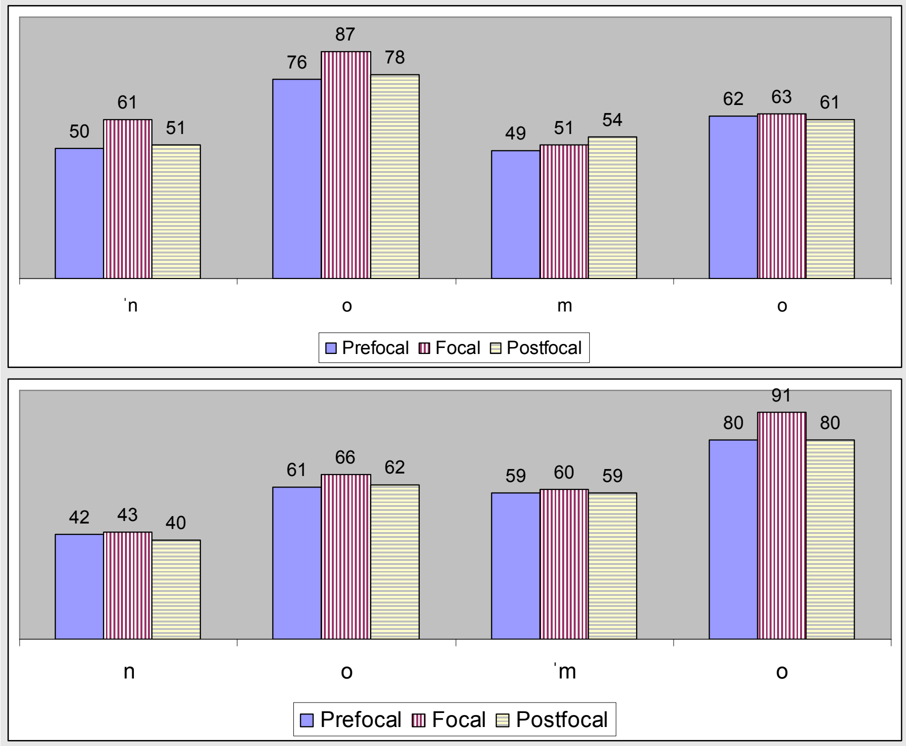


Figure 2.2. Five speakers' average segmental durations (grand mean) in ms of the minimal pair [ˈnomo~noˈmo] in the carrier sentence [i maˈria ˈiksere to ___ kaˈla] in prefocal, focal and postfocal position.

11 ms for the nasal (50~61 ms, t(4)=-2.4, p<0.05) and 11 ms for the vowel (76~87 ms, t(4)=-5.4, p<0.005); the unstressed syllables appeared with almost the same duration (111~114 ms). The duration differences of the second member of the stressed syllables between the pre(post)focal~focal executions were 12 ms (139~151 ms, t(4)=-3.3, p<0.025), the main contribution of the vowels (80~91 ms, t(4)=4.2, p<0.01), their nasals being fairly the same (59~60 ms); the unstressed syllable had a minor difference of 6 ms (103~109 ms, t(4)=-2.6, p<0.05), 5 ms for the vowels (61~66 ms, t(4)=-2.8, p<0.025), their nasals being of the same duration (42~43 ms).

Most of the speakers produced the majority of the stressed syllables and vowels in focus context with longer duration than the stressed syllables out of focus but not in a constant way. Speaker 2 produced the prefocal~focal stressed syllables of the second member of the pair with nearly the same duration (72-80~68-82 ms) whereas the syllabic duration between the focal and postfocal opposition appeared the same for speaker 3, the nasals and the vowels showing the opposite duration variations (58-90 ~ 66-82 ms).

To summarize the results of word and sentence stress duration pattern, the stressed syllables are longer than the unstressed ones and this distinction is mainly carried out by the vowels whereas the consonants[5] may even go in the opposite direction; this holds for both prefocal and postfocal positions. The stressed syllables in focus (sentence stress) are usually longer than out of focus (word stress) but not in a constant way and this distinction is mainly carried out by the vowels whereas the consonants may even for sentence stress go in the opposite direction. The unstressed syllables are minimally varied in different focal positions. Considering the prefocal and postfocal stressed~unstressed opposition as a whole with the unstressed syllables as reference, the stressed syllables are 24% longer than the unstressed ones out of focus (107~133 ms) and 39% longer when in focus (107~149 ms).

2.2.3.2 Fundamental Frequency

Figure 2.3 shows one speaker's raw F0 contours of the minimal pair [ˈnomo~noˈmo] in their carrier sentences with different contextual frames. We see that there is a partial but abrupt reorganization of the utterances with the application of the sentence stress, dividing the utterances into three parts: A prefocal part (lb~2b), which is correlated with a rise of F0 aligned with the production of stressed syllables; a focal part (lc~2c), which is correlated with an extended F0-rise followed by an abrupt F0-fall aligned with the production of sentence stress; and a postfocal part (ld~2d), which is correlated with a falling and flattening F0 to the end of the utterances.

[5]There is a tendency for the second nasal of the [ˈnomo~noˈmo] pair to appear longer than the first one, and this may depend on the intrinsic duration values that consonants with different places of articulation (dental~labial) may have in Greek.

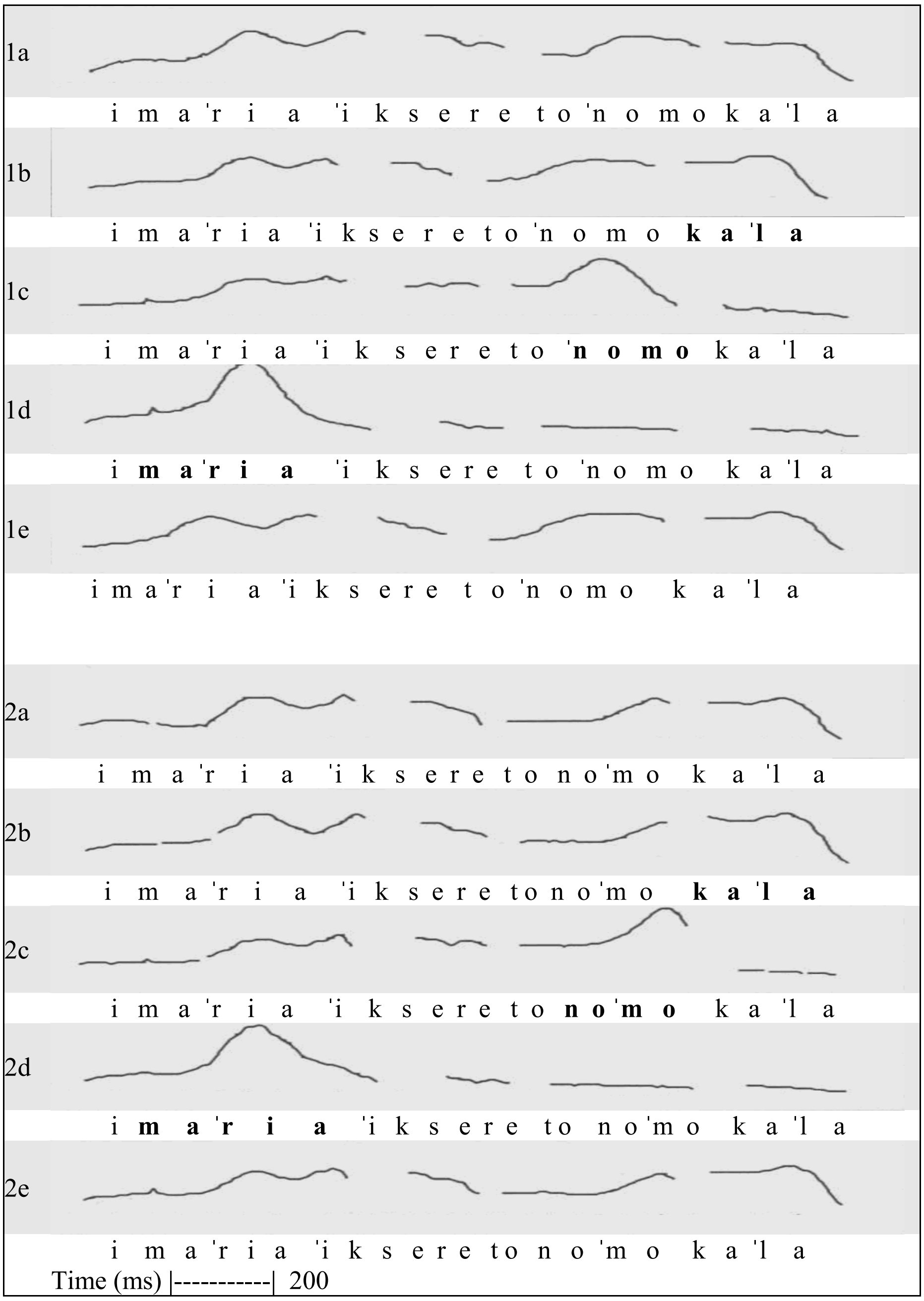


Figure 2.3. One speaker's raw F0 contours of the minimal pair [ˈnomo~noˈmo] (1~2) in the carrier sentence [i maˈria ˈiksere to ___ kaˈla] with different focal organizations (a-e).

Table 2.2. Five speakers' mean voice fundamental frequency (Hz), standard deviation (SD) and grand mean ($\bar{\bar{X}}$G) of the minimal pair [ˈnomo ~ noˈmo] in the carrier sentence [i maˈria ˈiksere to ___ **kaˈla]** in prefocal (lb~2b), focal (lc~2c) and postfocal position (ld~2d).

(lb~2b)	ˈn	o	m o	n	o	ˈm o
1 Hz	104	143	168	100	95	146
sd	10.8	10.9	11.5	6.1	7.0	24.0
2	130	173	180	131	122	178
	3.5	4.4	3.5	8.9	7.5	10.3
3	92	124	134	103	93	137
	5.7	11.4	8.2	5.7	6.7	4.4
4	137	174	183	131	121	170
	11.5	19.4	15.6	5.4	5.4	7.0
5	101	134	138	100	97	138
	4.1	4.1	4.4	2.6	2.1	3.2
$\bar{\bar{X}}$G	112	149	160	113	105	153
sd	19.5	22.8	23.1	16.4	14.5	18.9
(lc~2c)	**ˈn**	**o**	**m o**	**n**	**o**	**ˈm o**
1 Hz	133	242	81	130	140	250
sd	8.3	14.8	5.4	3.5	10.0	7.0
2	158	214	96	148	141	199
	8.3	13.4	3.6	10.3	11.4	12.4
3	129	164	78	132	130	172
	10.8	17.8	2.7	5.7	7.0	4.4
4	163	199	114	150	151	185
	4.4	8.9	5.4	3.5	2.2	5.0
5	108	156	86	102	102	163
	5.7	5.4	2.3	5.7	5.7	2.7
$\bar{\bar{X}}$G	138	195	91	132	132	193
sd	22.5	35.5	14.5	19.2	18.7	34.2
(ld~2d)	ˈn	o	m o	n	o	ˈm o
1 Hz	74	72	70	74	72	70
sd	2.4	2.4	2.4	1.6	1.6	1.6
2	99	97	89	95	94	92
	1.0	3.2	0.8	3.2	3.0	3.0
3	81	78	76	82	81	78
	2.2	1.7	1.7	1.4	1.0	1.0
4	116	117	109	110	109	106
	8.2	5.7	5.3	2.6	5.7	3.6
5	91	88	84	89	88	86
	2.6	1.7	1.5	6.9	7.7	8.2
$\bar{\bar{X}}$G	92	90	85	90	88	86
sd	16.3	17.6	14.9	13.6	13.9	13.7

Prefocal distribution

In prefocal position (Table 2.2, lb~2b), on the syntagmatic plane, the five speaker's mean F0 contour of the minimal pair [ˈnomo~noˈmo] was 112 Hz at the beginning of the nasal and 149 Hz at the end of the stressed vowel of the first member of the pair, a rise of 37 Hz; the F0 contour continued rising for 11 Hz, from 149 to 160 Hz, to the end of the unstressed syllable. At the second member of the pair, the F0 contour was falling for 8 Hz, from 113 to 105 Hz, during the unstressed syllable and then, at the beginning of the stressed syllable, it started rising from 105 to 153 Hz at the end of the syllable, a rise of 48 Hz. On the paradigmatic plane, the main F0 difference of the [ˈnomo~noˈmo] pair was at the syllable boundaries where F0 for [ˈnomo] was 149 Hz and for [noˈmo] 105 Hz, a difference of 44 Hz.

All five speakers produced both members of the pair with an F0 rise starting at the beginning of the stressed syllable (the nasal). This F0 rise aligned with the stressed syllable forms part of a larger F0 gesture which stretches to the whole stress group. Thus, although the stressed syllables of both members of the pair reach about the same maximum value, the post-stressed syllable of the first member is even higher whereas the pre-stressed syllable of the second member has a falling F0 contour as it is the last post-stressed syllable of the preceding stress group. The word stress F0 rise had a noticeable variation, some utterances were produced in a rather compressed form, particularly by speaker 1 for the second member (sd 24.0) and by speaker 4 for the first member of the pair (sd 19.4) in anticipation of the upcoming sentence stress.

Focal distribution

In focal position (Table 2.2, lc~2c), on the syntagmatic plane, the first member's syllable had an F0 rise of 57 Hz, from 138 Hz at the beginning of the nasal to 195 Hz at the end of the vowel; during the unstressed syllable there was an abrupt fall of 104 Hz, from 195 Hz at the beginning of the nasal to 91 Hz at the end of the vowel. For the second member of the pair, the unstressed syllable had a level F0 contour at 132 Hz; the stressed syllable had a variation of 61 Hz, from 132 Hz at the beginning of the nasal to 193 Hz in the second half of the vowel. On the paradigmatic plane, apart from the main difference of 63 Hz at the syllable boundaries (195~132 Hz) there was another difference of 102 Hz at the end of the words (91~193 Hz) as a result of the sentence stress abrupt fall of the first member.

All five speakers produced all utterances of both members of the pair in focus with an F0 rise from the beginning to the second half of the stressed syllables in combination with an abrupt F0 fall at the first post-stressed syllable rather than the whole stress group in a constant way. Moreover, all speakers reached the highest F0 point at the stressed syllable in focus for every single utterance.

Postfocal distribution

In postfocal position (Table 2.2, ld~2d), the F0 contour of the minimal pair [ˈnomo~noˈmo] was fairly the same, with no considerable variation between the stressed~unstressed syllables; it declined slightly from 92 to 85 Hz for the first member of the pair and from 90 to 86 Hz for the second member of the pair respectively.

Most of the speakers produced most of the utterances with no considerable F0 variation in relation to the stressed syllables in a rather constant way. The largest interspeaker variation appeared in speaker 4 who produced some of the utterances of the first member of the pair with minimal F0 rises (less than 5 Hz) aligned with the stressed syllables.

Contextual distribution

On the contextual plane (Figure 2.4), the first member's stressed syllable in focus started 26 Hz higher than the same syllable in prefocal position (138~112 Hz) and 46 Hz higher than the postfocal one (138~92 Hz), rose 46 Hz higher than the prefocal stressed syllable (195~149 Hz) and 105 Hz higher than the postfocal one(195~90 Hz); the unstressed syllable fell 69 Hz lower than the prefocal syllable (91~160 Hz) coming quite close to the values for the postfocal unstressed one (91~85 Hz). The second member's unstressed syllable in focus started 19 Hz higher than the prefocal syllable (132~113) and 42 Hz higher than the postfocal one (132~90 Hz), kept itself 27 Hz higher than the prefocal syllable (132~105 Hz) and 44 Hz than the postfocal one (132~88 Hz) to the end of the syllable; the stressed syllable in focus rose 40 Hz higher than the prefocal syllable (193~153 Hz) and 107 Hz higher than the post focal one (193~86 Hz).

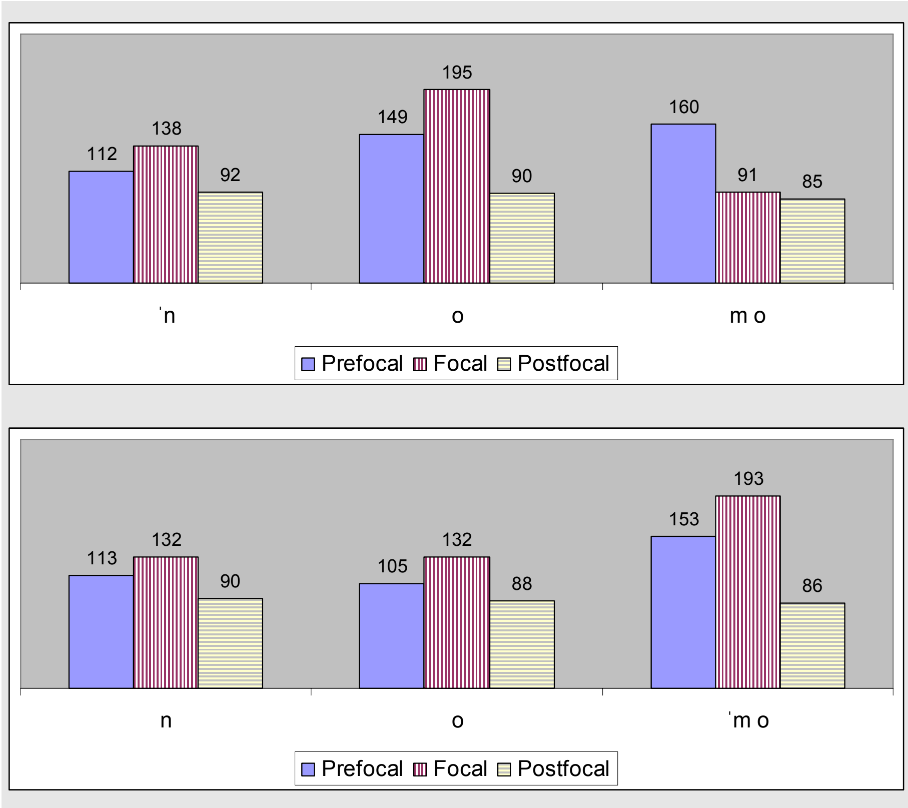


Figure 2.4. Five speakers' average F0 values (grand mean) in Hz of the minimal pair [ˈnomo~noˈmo] in the carrier sentence [i maˈria ˈiksere to ___ kaˈla] in prefocal, focal and postfocal position.

All five speakers produced the focal stressed syllables with the highest F0 top and the largest F0 fall into the first post-stressed syllable in a constant way across the utterances. The focal F0 rise was larger than the prefocal word stress rise but with a noticeable interspeaker variation; although speakers 1 and 5 produced both members of the pair with larger focal F0 rises, speaker 2 produced the second member with rather equal prefocal ~ focal rises of 56~58 Hz (122-178~141-199 Hz), speaker 3 produced both members with rather equal F0 rises of 32~35 Hz (92-124~129-164 Hz) for the first member and 44~42 Hz (93-137~130-172 Hz) for the second member of the pair, and speaker 4 had equal F0 rises for the first member, 37~36 Hz (137-174~163-199 Hz) as well as the greatest interspeaker variation for the second member of the pair where the prefocal F0 rise was

49 Hz (121-170 Hz) whereas the focal rise was 34 Hz (151-185 Hz). On the other hand, all speakers had the highest F0 contour at the beginning of the minimal pair in focus except speaker 4 whose prefocal-focal productions had little variation (100~102 Hz), as well as having the lowest postfocal F0 contour among the three focal productions.

To summarize the results of the F0 pattern of word and sentence stress, the prefocal word stressed syllables are aligned with an F0 rise whereas the focal ones, in addition to an F0 rise, are associated with a wide F0 range as well as an abrupt F0 fall; on the other hand, the postfocal word stressed syllables appear with minimal F0 variations. The word stress F0 rises are rather variable, the focal word stressed syllables may appear with the largest F0 rise within the utterance whereas the largest F0 range and the post-stressed abrupt fall are constant as well as the postfocal F0 variation which is minimal within the utterance. Furthermore, the word stress F0 changes are rather local at stress group level whereas the sentence stress has a global effect on the F0 pattern at utterance level.

2.2.3.3 Intensity

Figure 2.5 shows one speaker's raw intensity contours of the minimal pair [ˈnomo~noˈmo] in their carrier sentences in different contextual frames. In the majority of the utterances, the intensity contour is higher for the word stressed syllables than for the unstressed ones whereas for the sentence stressed syllables intensity is even higher, particularly on the paradigmatic plane. In prefocal position, production lb~2b, the stressed syllables have somewhat higher intensity than the unstressed ones. In focal position, productions lc~2c, the intensity contour of the stressed ~ unstressed opposition is enlarged as well as the whole intensity contour being somewhat raised. The postfocal syllables keep their stress ~ unstressed distinction at a lowered intensity contour (ld~2d), although for these particular productions only on the paradigmatic plane.

Prefocal distribution

In prefocal position (Table 2.3, lb~2b), on the syntagmatic plane, the five speakers' intensity contour of the minimal pair [ˈnomo~noˈmo] was quite even for the two syllables of the first member, (-10.3~-10.7 dB), whereas the second pair showed a variation of 4.2 dB, -14.1 dB for the unstressed syllable and -9.9 dB for the stressed syllable ($t(4)=-3.2$, $p<0.025$). On the paradigmatic plane, there was a significant difference of 3.8 dB between the first stressed~unstressed syllables (-10.3~-14.1 dB, $t(4)=2.7$, $p<0.025$) whereas the second unstressed~stressed syllables showed a negligible variation of 0.8 dB (-10.7~-9.9 dB).

All five speakers produced the majority of the stressed syllables of the second member of the pair with a higher intensity contour than its unstressed syllable. For the first member, speakers 1, 2 and 5 had higher intensity for the stressed syllables, speaker 3 equal intensity (-14.0~-14.0 dB) whereas speaker 4 showed the largest interspeaker variation with the stressed syllable having 1.8 dB lower intensity than the unstressed syllable (-8.2~-6.4 dB).

Focal distribution

In focal position (Table 2.3, lc~2c), on the syntagmatic plane, the stressed syllable of the first member was -6.7 dB whereas its unstressed syllable was -14.4 dB, a significant difference of 7.7 dB ($t(4)=6.2$, $p<0.005$); the unstressed syllable of the second member was -12 dB whereas its stressed syllable was -5.5 dB, a difference of 6.5 dB ($t(4)=-8.0$, $p<0.005$). The intensity distinction was kept on the paradigmatic plane as well, the first stressed~unstressed syllables having a difference of 5.3 dB (-6.7~-12 dB, $t(4)=4.9$, $p<0.005$) whereas the second unstressed~stressed syllables had a difference of 8.9 dB (-14.4~-5.5 dB, $t(4)=-7.0$, $\rho<0.005$).

All five speakers produced the stressed syllables in focus with higher intensity than the unstressed syllables in a rather constant way. The absolute top intensity values varied among the speakers as well as the stressed~ unstressed difference whereas the interspeaker difference variation between the members of the pair was rather limited, speaker 1 showing the largest variation, a difference of 11 dB (-1.6~-12.6 dB) for the first member and of 5 dB (-9.6~-4.2 dB) for the second member of the pair.

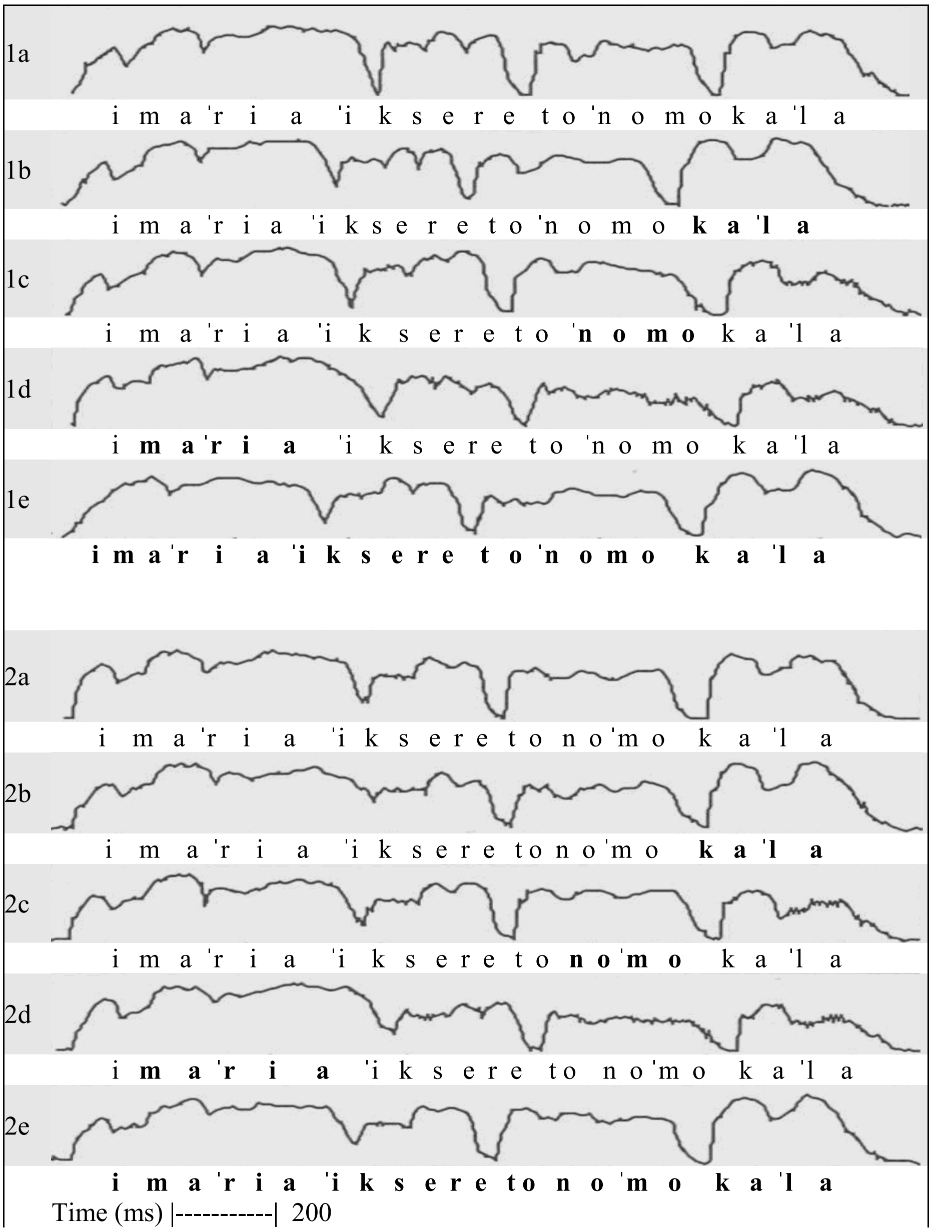


Figure 2.5. One speaker's raw intensity contours of the minimal pair [ˈnomo~noˈmo] (1~2) in the carrier sentence [i maˈria ˈiksere to ___ kaˈla] with different focal organizations (a-e).

Table 2.3. Five speakers' mean intensity (dB), standard deviation (sd) and grand mean ($\bar{\bar{X}}$G) of the niinimal pair [ˈnomo~noˈmo] in the carrier sentence [i maˈria ˈiksere to ___ kaˈla] in prefocal (lb~2b), focal (lc~2c) and postfocal (ld~2d) position.

(lb~2b)		ˈn o	m o	n o	ˈm o
1	dB	-7.6	r8.4	-14.0	-9.2
	sd	1.8	0.8	3.0	3.9
2		-11.8	-12.8	-14.0	-9.6
		2.0	1.6	2.0	3.0
3		-14.0	-14.0	-15.6	-14.2
		0.7	2.5	1.9	1.6
4		-8.2	-6.4	-15.8	-7.2
		2.8	1.1	1.6	1.7
5		-10.2	-12.2	-11.2	-9.4
		1.7	0.4	1.0	1.9
	$\bar{\bar{X}}$G	-10.3	-10.7	-14.1	-9.9
	sd	2.6	3.2	1.8	2.5
(lc~2c)		**ˈn o**	**m o**	**n o**	**ˈm o**
1	dB	-1.6	-12.6	-9.6	-4.2
	sd	0.5	2.7	0.8	2.5
2		-7.4	-15.0	-14.4	-6.2
		1.8	1.4	1.3	2.9
3		-12.6	-19.2	-14.8	-9.2
		0.5	2.2	0.8	1.0
4		-4.6	-14.0	-10.0	-1.4
		2.5	2.9	1.4	0.5
5		-7.6	-11.4	-11.2	-6.6
		1.8	0.8	1.0	1.1
	$\bar{\bar{X}}$G	-6.7	-14.4	-12.0	-5.5
	sd	4.0	2.9	2.4	2.9
(ld~2d)		ˈn o	m o	n o	ˈm o
1	dB	-17.6	-21.4	-18.8	-19.2
	sd	2.3	3.3	2.7	3.1
2		-16.2	-19.8	-20.4	-16.4
		0.8	1.0	0.8	1.3
3		-16.0	-19.6	-17.2	-13.2
		2.0	2.7	1.0	2.1
4		-12.6	-14.6	-16.8	-12.6
		1.9	1.5	2.0	1.3
5		-12.6	-17.2	-13.4	-11.6
		0.5	1.9	2.0	3.0
	$\bar{\bar{X}}$G	-15.0	-18.5	-17.3	-14.6
	sd	2.2	2.6	2.6	3.1

Postfocal distribution

In postfocal position (Table 2.3, ld~2d), on the syntagmatic plane, the stressed~unstressed syllables of the first member of the pair had a significant difference of 3.5 dB, (-15~-18.5 dB, $t(4)=8.3$, $p<0.005$); the unstressed~ stressed syllables of the second member of the pair had a difference of 2.7 dB (-17.3~-14.6 dB, $t(4)=-3.0$, $p<0.025$). On the paradigmatic plane, the first stressed~unstressed opposition had a significant difference of 2.3 dB, (-15~-17.3 dB, $t(4)=3.0$, $p<0.025$), whereas the second unstressed~stressed one had a significant difference of 3.9 dB (-18.5~-14.6 dB, $t(4)=-4.3$, $p<0.01$).

All five speakers produced the majority of the stressed syllables of both members of the pair with higher intensity than the unstressed syllables except speaker 1 who produced both syllables of the second member with almost equal intensity (-18.8~-19.2 dB).

Contextual distribution

On the contextual plane (Figure 2.6), both stressed and unstressed syllables of the first as well as the second member of the [ˈnomo~noˈmo] pair had significant variations for all three contextual oppositions. The first member's stressed syllable of the prefocal ~ focal opposition had a significant difference of 3.6 dB (-10.3~-6.7 dB, $t(4)=-4.6$, $p<0.005$), the focal ~ postfocal opposition 8.3 dB (-6.7~-15 dB, $t(4)=3.7$, $p<0.01$), and the prefocal ~ postfocal one a significant variation of 4.7 dB (-10.3~-15.0 dB, $t(4)=3.2$, $p<0.025$); the unstressed syllable of the prefocal ~ focal opposition had a significant variation of 3.7 dB (-10.7~-14.4 dB, $t(4)=2.5$, $p<0.05$), the focal ~ postfocal 4.1 dB (-14.4~-18.5 dB, $t(4)=2.5$, $p<0.05$) and the prefocal ~ postfocal 7.8 dB (-10.7~-18.5 dB, $t(4)=5.4$, $p<0.005$).

The second member's unstressed syllable of the prefocal~focal opposition showed a variation of 2.1 dB (-14.1~-12 dB, $t(4)=-1.6$, $p>0.05$), the focal~ postfocal 5.3 dB (-12~-17.3 dB, $t(4)=3.9$, $p<0.01$) and the prefocal~ postfocal 3.2 dB (-14.1~-17.3 dB, $t(4)=3.1$, $p<0.025$); the stressed syllable of the prefocal~focal opposition had a variation of 4.4 dB (-9.9~-5.5 dB, $t(4)=-7.8$, $p<0.005$), the focal ~ postfocal 9.1 dB (-5.5~-14.6 dB, $t(4)=4.4$, $p<0.01$) and the prefocal ~ postfocal 4.7 dB (-9.9~-14.6 dB, $t(4)=2.4$, $p<0.05$).

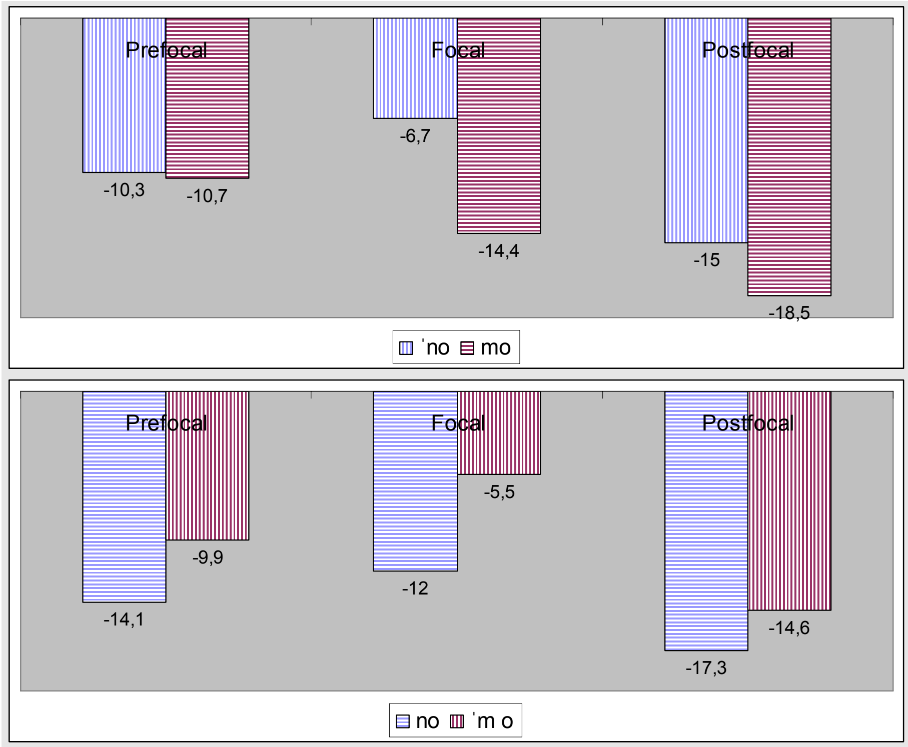


Figure 2.6. Five speakers' average intensity (grand mean) in dB of the minimal pair [ˈnomo~noˈmo] in the carrier sentence [i maˈria ˈiksere to ___ kaˈla] in prefocal, focal and post focal position.

All five speakers produced the stressed syllables in focus with higher intensity than out of focus of both members of the pair in a constant way. The difference between the stressed~unstressed syllables in focus was also larger than out of focus for both members of the pair except for speaker 1 who produced the prefocal, second member with a difference of 4.8 dB (-14.0~-9.2 dB) as well as the focal, second member with an almost equal difference of 5.4 dB (-9.6~-4.2 dB).

To summarize the results of the intensity pattern of word and sentence stress, the word stress distinction is produced with a rather regular higher intensity whereas sentence stress is realized with even higher absolute intensity as well as larger intensity differences between the stressed and unstressed syllables in a fairly constant way.

2.2.4 Discussion

To recapitulate the results of this experiment, word stress is realized by the contribution of duration, F0, and intensity before focus and by a combination of duration and intensity after focus; sentence stress is realized as an extended increase of duration, F0, and intensity locally and as a reorganization of the prosodic structure globally. The results of this experiment quantitatively corroborate our earlier findings of an investigation on the contribution of the acoustic parameters of duration, F0, and intensity to the realization of word and sentence stress in different sentence positions and contextual frames in Greek (Botinis 1982).

Our approach is based on the fundamental concept that different prosodic contributions may merge into the same speech unit (Bruce 1977) and our aim has been to isolate the acoustic parameters in contexts where the influence of different prosodic categories is minimal. Furthermore, the contribution of the acoustic parameters to different prosodic categories as well as to the same one may vary, i.e. we assume a basic independence of the prosodic features (Botinis 1982), although they may be interrelated (Bannert 1986).

In prefocal position, the word stressed syllables are longer than the unstressed ones, and this distinction is mainly carried out by the vowels. The distinction is even kept when the stressed syllable is in a non-final position versus a final unstressed position, i.e. stress has a stronger effect on vowel duration than the probable phrase final lengthening (Lindblom and Rapp 1973, Klatt 1976). Greek does not have a phonological distinction of length, therefore duration should be a strong candidate for stress distinction according to Berinstein (1979).

Duration has been found to be correlated with stress in a number of languages like Danish (Fischer-Jørgensen 1984), English (Fry 1958), German (Bannert 1985), Italian (Farnetani and Kori 1986), Swedish (Strangert 1985), among others. That does not necessarily mean that duration is a phonetic universal of stress since for languages like Hungarian (Fonage 1966) or Welsh (Williams 1986) stressed syllables may be shorter than unstressed syllables.

F0 starts rising at the beginning of the stressed syllable and reaches the highest point at the end of it or the first post-stressed syllable and then falls gradually to the end of the stress group irrespective of word or phrase boundaries; by stress group we mean the stressed syllable and any unstressed syllable(s) up to the next stressed syllable (Thorsen 1982). Usually a stress group is associated with a tonal gesture although the tonal gesture may be assimilated to a larger, focal gesture in which case the stress group boundaries function as fixed points of coordination between rhythm and intonation.

The importance of the F0 turning-points has been emphasized by Gårding (1977b), and an F0 change may be interrupted or even omitted by the influence of other prosodic categories in Swedish, but never displaced in time (Bruce 1983). Thus, in Greek, the F0 variations associated with word stress may be flattened up to focus and, in such cases, the F0 contour is realized as a continuously rising F0, i.e. we have a compressed F0 manifestation (Gårding 1984); similar F0 manifestations in prefocal position have been reported for German (Bannert 1985).

The unmarked realization of word stress in prefocal position in Greek is an F0 rise (Botinis 1982). In Standard Swedish the unmarked realization of the word accent is an F0 fall in relation to the stressed syllable, earlier for the acute accent and later for the grave accent (Bruce 1977), although the word accent distinction may be either an F0 fall or an F0 rise in certain context dependent conditions (Bruce 1983). Hyman (1977) reports that in several languages stress may be realized both as a rising and a falling F0; this has been observed for Dutch ('t Hart and Cohen 1973), English (Bolinger 1958, O' Shaughnessy 1979), Swedish (Gårding et al. 1970, Hadding-Kock 1961), to name some.

The peak intensity of the stressed syllables is higher than the unstressed ones on the syntagmatic and/or paradigmatic plane. In Greek, intensity has been found highly correlated with duration but this is not necessarily always the case (Botinis 1982). The close relation between intensity and duration to denote stress has urged some phoneticians to take into account the “energy integral” and not only the intensity peak, combining thus both intensity and duration into a single feature (Lieberman 1960, Gårding 1967, Lea 1977, Fischer-Jørgensen 1984, Beckman 1986).

On the other hand, intensity and F0 have been reported interdependent and highly correlated (Lieberman 1967, Ohala 1978); this may be the unmarked realization of word stress in neutral utterances. But as soon as an element of the utterance is focused the prosodic parameters are restructured and we may have an F0 rise combined with an intensity fall in prefocal position in Greek, and not only at the end of the interrogative sentences in English (Lieberman 1967). Hirst (1981) has found intensity and F0 correlated during the first part of the vowel /a/ in French but the opposite in the second part of it and remarks the absence of a total correlation between the two parameters.

In focal position, the stressed syllables are even longer than when they are out of focus and the duration contribution of the sentence stress is mainly carried by the vowels. Bruce (1981) found the consonants mainly affected by the focal lengthening in Swedish and explains the phenomenon by the fact that in his speech material the affected syllable has had a phonologically long consonant which is in accordance with the hypothesis that it is the phonologically long segment of the syllable which is most greatly increased by sentence stress (Bannert 1979).

In the present experiment, the majority of the speakers produced the utterances with longer duration in the stressed syllable of the word in focus but not in a constant way; some productions appeared with the same duration in focal and non-focal position. In our (1982) study, the two speakers hardly showed any duration contribution to the production of sentence stress. Our combined results show that duration may not be a constant acoustic parameter for sentence stress. Similar results with interspeaker variation in using duration for focus have been reported for Danish (Thorsen 1980), German (Bannert 1979), and Swedish (Bruce 1981) whereas in French (Touati 1987) the contribution of duration to focal accent (accent d'insistance) in non final position is minimal.

F0 manifestation for focus has three dimensions. First, an enlarged F0 range followed by an abrupt F0 fall associated with the sentence stress; second, an optional F0 compression for any prefocal word stress(es) and, third, an obligatory F0 fall and flattening for any postfocal speech material. The F0 structure across the Greek utterances may be properly described along Gårding's (1984) later version model in which the F0 contour may appear in an expanded, compressed, or one line form. An extended F0 rise for sentence stress has been observed in Danish (Thorsen 1980), English (O' Saughnessy 1979), French (Touati 1987), German (Bannert 1985), Greek (Botinis 1982), and Swedish (Bruce 1977), among others.

Apart from the global manifestation of focus in Greek, a reorganisation of the whole utterance (Rossi 1985), the local effects of sentence stress is an extended F0 range followed by an abrupt F0 fall irrespective of word, phrase, or stress group boundaries. In standard Swedish, the sentence accent is realized as a post-word accent F0 rise (Bruce 1977) whose domain is the whole stress group (Bruce 1983); in the southern Swedish dialect, the sentence accent is realized as an extended F0 fall on the word stressed syllable. It seems obvious from the afore mentioned examples that different languages (Gårding et al. 1982), and even different dialects (Gårding and Bruce 1981), may vary the way they distribute the same acoustic parameter, F0, to denote a corresponding prosodic category like word or sentence stress.

The intensity of the stressed syllables in focus is even higher than out of focus. Sentence stress in Greek is usually realized with longer duration, higher F0, and higher intensity; this by no means implies that these acoustic parameters are dependent on each other. The fact that they are combined in different ways across the utterance points to their independence (Botinis 1982). Since focus is related to the most important unit of information, the three acoustic parameters seem to contribute synergetically (Bannert 1982), though intensity and F0 appear as the most preferable ones across subjects.

Beckman (1986) reports the energy integral (total amplitude) as the main acoustic correlate of nuclear stress in English. The Greek data supports the correlation between energy integral and word stress out of focus (see below). In focal position (sentence stress, nuclear stress) duration is somewhat variable whereas F0 has been found to be highly constant among the speakers. This discrepancy between Beckman's and our results may be partly due to the different experimental setups used to produce nuclear and sentence stress respectively.

In postfocal position, the word stressed syllables are longer than the unstressed ones, the stressed~unstressed opposition being very similar to the prefocal one. Our results are in agreement with Fourakis' (1986) who found that the stressed syllables were 26.4% longer than the unstressed ones in slow tempo and 25.4% longer in fast tempo in Greek. Nakatani and Aston, reported in Beckman (1986, p. 60), found duration contribution in postnuclear position in English whereas Huss (1978) attributes the postnuclear variation of duration to rhythmic structuring rather than to the production of lexical stress.

The amount of duration used in the production of stress seems language specific. Thus, in English, Klatt (1976) reports that the average duration of a stressed vowel in connected discourse is about 130 ms whereas the unstressed vowels as well as the consonants are about 70 ms long, i.e. a lengthening of about 42% of the stressed syllable. In Swedish (Strangert

1985), which has a phonological length distinction, the stressed syllables are twice as long as the unstressed ones whereas in Danish (Fischer-Jørgensen 1984), a Scandinavian language with length distinction, too, the stressed~unstressed ratio is smaller. On the other hand, in Spanish, reported in Strangert (1985), the stressed syllables are 20% longer than the unstressed ones whereas in Italian (Farnetani and Kori 1986) the difference is larger.

The F0 of the postfocal syllables declines at a low level with hardly any variation for the stressed~unstressed distinction, what Gårding (1984) describes as a falling, one-line contour. The postfocal F0 flattening has been observed in Danish (Thorsen 1980), English (O' Shaughnessy 1979), Finland Swedish (Tevajärvi 1982), French (Touati 1987), German (Bannert 1985), Greek (Botinis 1982), among others.

In Standard Swedish (Bruce 1977), the word accent distinction is kept even after focus, although in a rather compressed form, as a down-stepping F0 contour (Bruce 1982). In Southern Swedish (Gårding et al. 1982), there is no postfocal reorganization of the F0-contour in contrast to Greek and French with the regular postfocal F0-flattening. Thus, although postfocal F0 flattening is quite common for a number of languages, it may vary between dialects (Standard~Southern Swedish), between closely related languages (Danish~Swedish), and be the same between more loosely related languages like Danish, French, German, and Greek.

The intensity contour of the postfocal word stressed syllables is higher than the unstressed ones (unlike the F0 contour) which combined with duration forms the energy integral, a fairly constant acoustic correlate in this position. The use of different combinations of the acoustic parameters to denote the same prosodic category, word or sentence stress, strongly indicate the independent nature of these parameters (Hirst 1981, Bannert 1986) as well as their inter and intraspeaker variation which reveals a synergetic (Bannert 1982) rather than an absolute contribution.

A hierarchy of the acoustic cues to describe stress (Fry 1958, Berinstein 1979) appears too simple since the acoustic parameters contribute in different ways across the syntagm. It seems, however, that the energy integral has been lexicalized in Greek whereas the voice fundamental frequency functions at a higher level attributing the sentence elements the appropriate semantic weighting (focus) in accordance with the discourse requirements (see part 4).

2.3 Acoustic Parameters of Enclitic Stress

2.3.0 Introduction

The main purpose of the second acoustic experiment is to investigate the contribution of duration, F0, and intensity to the production of enclitic stress in prefocal, focal and postfocal position. Enclitic stress is the prosodic category whose domain is the phrase, its function is to contribute to rhythmic structuring in combination with word stress, and it is morpholexically and syntactically determined. Moreover, the enclitic structure with a word and an enclitic stress was compared with the proclitic structure with a word stress.

In this experiment the following questions are addressed: (1) What is the acoustic structure of enclitic stress? (2) Is the acoustic structure of enclitic stress the same as the one of word stress or sentence stress, or does enclitic stress form an acoustic structure of its own? (3) Apart from enclitic stress, does the enclitic structure differ in any other respect from the proclitic one?

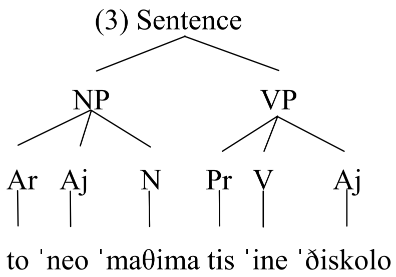
In establishing prosodic categories we assume the following taxonomic criteria: every prosodic category has (i) its distribution rules at the respective level, (ii) its phonetic realization in comparison to other prosodic categories, and (iii) its distinct function in the linguistic system of the language.

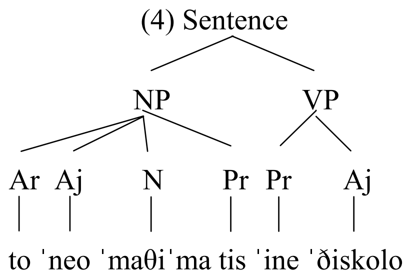
In the first acoustic experiment we have separated word from sentence stress, i.e. apart from their distinct distribution and function (part 1) word stress is realized by variations of duration and intensity, i.e. the energy integral, whereas sentence stress by variations of voice fundamental frequency. The separation of word and sentence stress has been a question with regards to their nature for the last decades. The American tradition (Trager and Smith 1951, Chomsky and Halle 1968, Liberman and Prince 1977), considers stress and pitch as independent prosodic features, although sentence stress often coincides with the intonation center of an utterance. The European tradition (Danes 1967, Halliday 1967, Crystal 1969, Bruce 1977, Rossi 1985), on the other hand, considers sentence stress as an intonation feature which coincides with a stressed syllable of an utterance.

> For the Americans, sentence stress is primarily a stress phenomenon which is often associated with a pitch change, while for the British it is an intonational phenomenon which occurs at a stressed syllable. The disagreement between the two traditions concerns the nature of sentence stress, not the separability of stress and intonation. (Ladd 1980, p. 8)

2.3.1 Speech Material

Two meaningful Greek sentences were set up; the first sentence with the proclitic structure, [to ˈneo ˈmaθima tis ˈine ˈðiskolo] (the new lesson is difficult for her), and the second one with the enclitic structure, [to ˈneo ˈmaθiˈma tis ˈine ˈðiskolo] (the new lesson of hers is difficult). The criteria for the test sentences were naturalness and simplicity of the sentences rather than a selection of words composed by sonorants with an unbroken F0 contour; yet, the word [ˈmaθima] has been chosen with identical sonorants in the first and third syllable where the word and enclitic stress are manifested. Both sentences are declarative composed of a noun phrase and a verb phrase though with different syntactic structures according to (3) and (4).





Each of the test sentences is the carrier of five different contextual information frames listed in Table II.II. The five productions were elicited as answers formulated in different ways to make the speaker choose one of the elements of the sentence as the focus and the carrier of the information required by the questioner. In productions 3a~4a the adjective [ˈðiskolo] was in focus and the minimal structure [ˈmaθima tis~ˈmaθiˈma tis] in prefocal position. In 3b~4b the noun [ˈmaθima] as well as the proclitic~enclitic [tis] was in focus and in 3c~4c only the noun [ˈmaθima] was in focus and the proclitic~enclitic [tis] out of focus. In 3d~4d only the enclitic~proclitic [tis] was in focus; although the clitic [tis] is a pronoun and in everyday language it should be represented by the noun it stands for, speakers properly instructed had no difficulties in applying focus to it for the purpose of the experiment. Last, in 3e~4e the adjective [ˈneo] was in focus and [ˈmaθima tis~ˈmaθiˈma tis] in postfocal position.

Table II.II. The minimal pair [ˈmaθima tis~ˈmaθiˈma tis] in the carrier sentence [to ˈneo ___ ˈine ˈðiskolo] with different contextual organizations (see text).

Contextual frame	Test sentences
3a. [ti tis ˈine to ˈneo ˈmaθima] (What is the new lesson for her?)	[to ˈneo ˈmaθima tis ˈine ˈ**ðiskolo**] (The new lesson is **difficult** for her.)
3b [pço ˈneo ˈine ˈðiskolo] (Which new is difficult?)	[to ˈneo ˈ**maθima tis** ˈine ˈðískolo] (The new **lesson** is difficult for **her**.)
3c. [pço ˈneo tis ˈine ˈðiskolo] (Which new is difficult for her?)	[to ˈneo ˈ**maθima** tis ˈine ˈðiskolo] (The new **lesson** is difficult for her.)
3d. [ja pçon ˈine ˈðiskolo to ˈneo ˈmaθima] (For whom is the new lesson difficult?	[to ˈneo ˈmaθima **tis** ˈine ˈðiskolo] (The new lesson is difficult for **her**.)
3e. [pço ˈmaθima tis ˈine ˈðiskolo] (Which lesson is difficult for her?)	[to ˈ**neo** ˈmaθima tis íne ˈðiskolo] (The **new** lesson is difficult for her.)
4a. [ti ˈine to ˈneo ˈmaθiˈma tis] (What is the new lesson of hers?)	[to ˈneo ˈmaθiˈma tis ˈine ˈ**ðiskolo**] (The new lesson of hers is **difficult**.)
4b. [pço ˈneo ˈine ˈðiskolo] (Which new is difficult?)	[to ˈneo ˈ**maθiˈma tis** ˈine ˈðiskolo] (The new **lesson of hers** is difficult.)
4c. [ˈpço ˈneo tis ˈine ˈðiskolo] (Which new of hers is difficult?)	[to ˈneo ˈ**maθiˈma** tis ˈine ˈðiskolo] (The new **lesson** of hers is difficult.)
4d. [pçaˈnu to ˈneo ˈmaθima ˈine ˈðiskolo] (Whose new lesson is difficult?)	[to ˈneo ˈmaθiˈma **tis** ˈine ˈðiskolo] (The new lesson of **hers** is difficult.)
4e. [pço ˈmaθiˈma tis ˈine ˈðiskolo] (Which lesson of hers is difficult?)	[to ˈ**neo** ˈmaθiˈma tis ˈine ˈðiskolo] (The **new** lesson of hers is difficult.)

2.3.2 Results

The set up of the material in this experiment permits comparisons of the proclitic~enclitic structure on all three planes, syntagmatic ~ paradigmatic ~ contextual. On the syntagmatic plane, the proclitic structure's antepenultimate syllable of the word [ˈmaθima] carrying word stress can be compared with the ultimate, unstressed syllable, a subject covered, by and large, in the first experiment. On the other hand, the enclitic structure's antepenultimate syllable of the word [ˈmaθiˈma] carrying word stress can be compared with the ultimate, enclitic stressed syllable.

On the paradigmatic plane, the antepenultimate, word stressed syllable as well as the ultimate, unstressed syllable of the proclitic structure can be compared with the antepenultimate, word stressed syllable and the ultimate, enclitic stressed syllable of the enclitic structure respectively.

On the contextual plane, the proclitic structure carrying word stress as well as the enclitic structure carrying both word and enclitic stress with different elements in focus is compared in prefocal, focal, and postfocal position.

As an example, one speaker's raw data of duration, F0, and intensity of the minimal structure [ˈmaθiˈma tis ~ ˈmaθiˈma tis] in different contextual frames is shown in Figures 2.7, 2.9, 2.11; the five speakers' mean values are shown in Tables 2.4-2.6 and Appendices 2.1-2.3, their grand mean in Figures 2.8, 2.10, 2.12.

From a first inspection of the material (Figure 2.7) we see that different focus applications are neutralized with the distribution of sentence stress. Thus, the proclitic structure [ˈmaθima tis] has the same prosodic manifestation whether both the word and the proclitic or only the word is in focus (3b~3c); on the other hand, when the proclitic [tis] is in focus, the prosodic manifestation of both the word and the proclitic is the same as with the prefocal manifestation (3d~3a), i.e. sentence stress is neither applied to the focal element nor to the proclitic's prosodic host but rather to its syntactic host. On the contrary, when the enclitic structure [ˈmaθiˈma tis] is in focus, either as a whole (4b) or any element of it (4c~4d), the prosodic manifestation of the phrase is quite the same. Thus, the basic prefocal, focal, and postfocal classification of the material holds even when enclitic stress is involved, although sentence stress realization may not coincide with the focal element.

In the following presentation, we shall concentrate on 3a, 3b, 3e ~ 4a, 4b, 4e, i.e. the proclitic~enclitic structure in prefocal, focal and postfocal positions, and leave out 3c as being equal in prosodic manifestation to 3b, 3d to 3a, and 4d, 4c to 4b.

2.3.2.1 Duration

In Figure 2.7, one speaker's durations of the minimal structure [ˈmaθima tis ~ ˈmaθiˈma tis] in different contextual frames is shown. Although there is a temporal reorganization of the entire utterances as a result of the application of sentence stress, the duration pattern of the prosodic categories seems quite constant: Word stressed syllables are longer than unstressed syllables (3a, 3e), enclitic stressed syllables are longer than word stressed syllables (4a, 4e) and sentence stressed syllables are longer than both word and enclitic stressed syllables regardless of whether sentence stress is realized on syllables already having word (3b) or enclitic stress (4b).

Prefocal distribution

In prefocal position (Table 2.4, 3a~4a), on the syntagmatic plane, the word stressed syllable and its segments were significantly longer than the unstressed syllable of the first member of the [ˈmaθima tis ~ ˈmaθiˈma tis] pair. The word stressed syllable was 43 ms longer than the unstressed one (160~117 ms, $t(4)=6.5$, $p<0.005$), 12 ms longer for the nasal (62-50 ms, $t(4)=3.2$, $p<0.025$) and 31 ms longer for the vowel (98~67 ms, $t(4)=4.8$, $p<0.005$). In the second member of the pair, the enclitic stressed syllable was 11 ms longer than the word stressed syllable (146-135 ms, $t(3)=0.6$, $p>0.05$); the nasals of the compared syllables were of the same duration (55 ms) whereas the enclitic stressed vowel was 11 ms longer than the word stressed one (91-80 ms, $t(3)=1.2$, $p>0.05$).

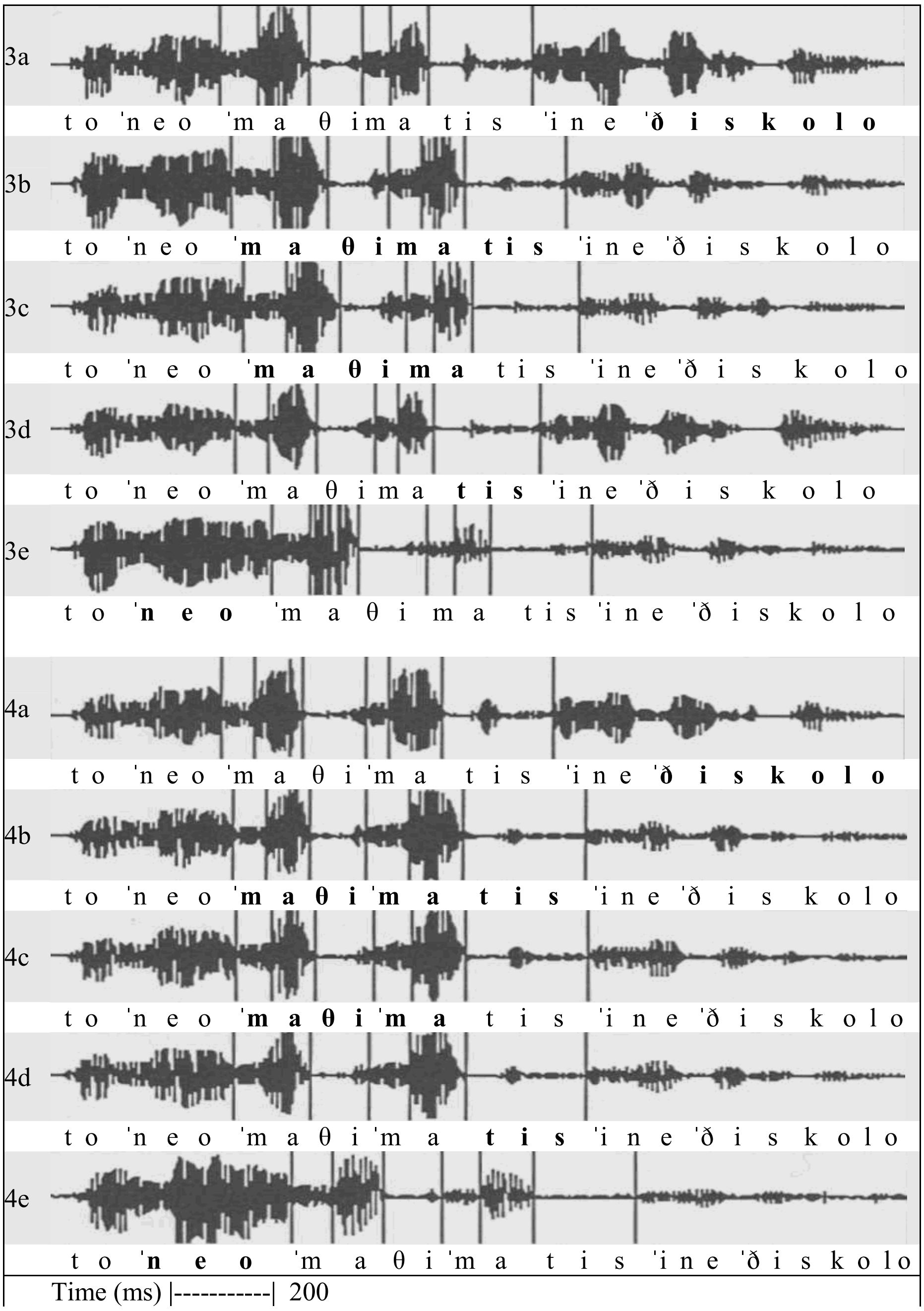


Figure 2.7. One speaker's raw durations of the minimal pair [ˈmaθima tis ~ ˈmaθiˈma tis] (3~4) in the carrier sentence [to ˈneo ___ ˈine ˈðiskolo] with different focal organizations (a-e).

Table 2.4, Five speakers' mean durations (ms), standard deviation (sd) and grand mean ($\bar{\bar{X}}$G) of the minimal pair [ˈmaθima tis~ˈmaθiˈma tis] in the context [to ˈneo ___ ˈine ˈðiskolo], in prefocal (3a~4a), focal (3b~4b), and postfocal position (3e~4e). *Speaker 4 is not included in the calculation of the grand mean ($\bar{\bar{X}}$G) of the enclitic structure in 3a~4a.

(3a~4a)		ˈm	a	θ i	m	a	tis	ˈm	a	θ i	ˈm	a	tis
1	ms	73	90	101	53	68	203	62	86	104	48	88	212
	sd	5.1	6.3	7.5	10.3	4.0	5.1	4.4	5.4	5.4	4.4	8.3	10.9
2		64	120	94	64	74	182	54	86	94	82	120	188
		8.9	7.0	5.4	11.4	5.4	8.3	5.4	5.4	5.4	4.4	7.0	13.0
3		64	104	94	44	56	188	56	82	86	48	92	172
		5.4	5.4	5.4	5.4	8.9	8.3	5.4	8.3	5.4	4.4	8.3	8.3
4*		54	104	82	44	84	186	(46	84	90	50	106	204)
		5.4	5.4	8.3	5.4	5.4	11.4	(5.4	5.4	10.0	7.0	5.4	5.4)
5		58	74	84	48	54	160	48	68	74	44	64	144
		4.4	5.4	5.4	4.4	5.4	7.0	4.4	4.4	5.4	5.4	5.4	5.4
	$\bar{\bar{X}}$G*	62	98	91	50	67	183	55	80	89	55	91	179
		7.1	17.2	7.8	8.3	12.5	15.4	5.7	8.5	12.6	17.7	22.9	28.5
3b~4b)		**ˈm**	**a**	**θ i**	**m**	**a**	**tis**	**ˈm**	**a**	**θ i**	**ˈm**	**a**	**tis**
1	ms	70	108	106	68	70	202	66	88	108	56	120	210
	sd	7.0	8.3	5.4	4.4	7.0	8.3	5.4	4.4	8.3	5.4	10.0	4.4
2		84	126	94	66	76	188	62	84	98	70	116	200
		5.4	5.4	5.4	5.4	5.4	8.3	4.4	5.4	8.3	7.0	5.4	12.2
3		68	118	96	62	58	188	50	86	90	64	96	194
		4.4	4.4	5.4	4.4	4.4	10.9	7.0	5.4	7.0	5.4	5.4	8.9
4		54	112	100	50	86	198	48	86	98	54	114	202
		5.4	4.4	7.0	7.0	8.9	8.3	4.4	5.4	4.4	5.4	5.4	8.3
5		56	94	86	44	58	166	34	84	84	44	82	160
		5.4	5.4	5.4	5.4	4.4	5.4	5.4	5.4	5.4	5.4	8.3	7.0
	$\bar{\bar{X}}$G	66	111	96	58	69	188	52	85	95	57	105	193
	sd	12.1	11.9	7.4	10.4	12.0	13.9	12.6	1.6	9.0	9.9	16.0	19.4
(3e~4e)		ˈm	a	θ i	m	a	tis	ˈm	a	θ i	ˈm	a	tis
1	ms	74	92	112	62	72	206	64	88	110	56	92	206
	sd	5.4	4.4	4.4	8.3	4.4	5.4	8.9	4.4	10.0	5.4	4.4	5.4
2		68	100	94	66	76	188	66	96	100	74	92	220
		4.4	7.0	5.4	5.4	5.4	8.3	5.4	5.4	7.0	5.4	4.4	12.2
3		54	82	96	50	64	200	50	74	94	66	92	192
		5.4	8.3	5.4	7.0	8.9	12.2	10.0	5.4	5.4	8.9	4.4	8.3
4		52	104	84	52	80	192	46	84	96	56	106	204
		4.4	8.9	5.4	8.3	7.0	10.9	5.4	8.9	5.4	5.4	5.4	11.4
5		56	76	80	48	56	158	44	66	84	44	68	156
		5.4	5.4	7.0	4.4	5.4	8.3	4.4	5.4	5.4	5.4	8.3	8.9
	$\bar{\bar{X}}$G	60	90	93	55	69	188	54	81	96	59	90	195
	sd	9.6	11.7	12.4	7.9	9.6	18.5	10.2	11.7	9.4	11.3	13.7	24.2

On the paradigmatic plane, the proclitic's word stressed syllable was significantly longer than the enclitic's word stressed syllable and its nasal. The proclitic's word stressed syllable was 25 ms longer than the enclitic's one (160~135, $t(3)=3.1$, $p<0.025$), 7 ms longer for the nasal (62~55 ms, $t(3)=15.4$, $p<0.0005$) and 18 ms longer for the vowel (98~80 ms, $t(3)=2.3$, $p>0.05$). The enclitic stressed syllable of the enclitic structure as well as its segments were significantly longer than the proclitic's unstressed ultimate as well as its segments. The enclitic structure's enclitic stressed syllable was 29 ms longer than the proclitic structure's unstressed syllable (146~117, $t(3)=2.4$, $p<0.05$), 5 ms longer for the nasal (55~50 ms, $t(3)=0.6$, $p>0.05$) and 24 ms longer for the vowel (91~67 ms, $t(3)=3.4$, $p<0.025$). The medial unstressed syllables showed a negligible difference of 2 ms (91-89) in favour of the proclitic structure whereas the proclitic/enclitic syllable [tis] showed a minor difference of 4 ms (183~179), also in favour of the proclitic structure.

All speakers produced the stressed syllable of the proclitic structure as well as its vowel with longer duration than its ultimate, unstressed syllable, in a constant way whereas the nasal did not show the same degree of constancy; speaker 2 produced the nasal of both syllables with the same duration (64 ms). The word and enclitic stress of the enclitic structure showed a high degree of variation among the speakers. Speaker 2 produced the enclitic stressed syllable with longer duration (202~140 ms), speaker 3 with about the same duration (140~138 ms), and speakers 1 (136~148 ms) and 5 (108-116 ms) with shorter duration than the word stressed syllable. The vowel of the enclitic stressed syllable was longer than the vowel of the word stressed syllable with the exception of speaker 5 (64~68 ms) whereas the word stress nasal was longer than the enclitic stress nasal except in the case of speaker 2 (54~82 ms).

On the other hand, all speakers produced both nasal and vowel of the word stressed syllable of the proclitic structure with longer duration than the nasal and vowel of the word stressed syllable of the enclitic structure with high constancy. The enclitic stressed syllable of the enclitic structure was longer than the ultimate, unstressed syllable, of the proclitic structure for all speakers, the vowel was constantly longer whereas the nasal was longer only for speakers 2 (82~64 ms) and 3 (48~44 ms).

Focal distribution

In focal position (Table 2.4, 3b~4b), on the syntagmatic plane, the proclitic's word stressed~unstressed syllables as well as their segments had significant differences. The word stressed syllable of the proclitic structure was 50 ms longer than the unstressed one (177~127 ms, $t(4)=6.8$, $p<0.005$), 8 ms longer for the nasal (66~58 ms, $t(4)=2.8$, $p<0.025$) and 42 ms longer for the vowel (111~69 ms, $t(4)=7.1$, $p<0.005$). The enclitic structure's enclitic stressed ~ word stressed syllables had a significant differences. The enclitic stressed syllable was 25 ms longer than the word stressed syllable (162~137 ms, $t(4)=4.6$, $p<0.05$), 5 ms longer for the nasal (57~52 ms, $t(4)=1.3$, $p>0.05$) and 20 ms longer for the vowel (105~85 ms, $t(4)=2.9$, $p<0.025$).

On the paradigmatic plane, the proclitic structure's word stressed syllable and its segments were significantly longer than the enclitic structure's ones. The word stressed syllable of the proclitic structure was 40 ms longer than the word stressed syllable of the enclitic structure (177~137 ms, $t(4)=5.5$, $p<0.005$), 14 ms longer for the nasal (66~52 ms, $t(4)=3.6$, $p<0.025$) and 26 ms longer for the vowel (111~85 ms, $t(4)=4.8$, $p<0.005$). The enclitic structure's enclitic stressed syllable was significantly longer than the proclitic structure's unstressed syllable. The enclitic stressed syllable was 35 ms longer than the unstressed syllable (162~127 ms, $t(4)=10.2$, $p<0.0005$); the contrasted nasals had no duration difference (57-58 ms) whereas the enclitic stressed vowel was (36 ms) significantly longer than the unstressed one (105~69 ms, $t(4)=7.8$, $p<0.005$). The medial unstressed syllables [θi] showed no difference between the proclitic/enclitic structure (96~95 ms) whereas the proclitic/enclitic [tis] showed a minor difference of 5 ms (188~193 ms, $t(4)=1.5$, $p>0.05$) in favour of the enclitic structure.

For all speakers both nasal and vowel of the word stressed syllable of the proclitic structure were longer than its ultimate, unstressed syllable, with high constancy. Both nasal and vowel of the enclitic stressed syllable of the enclitic structure were longer than the word stressed syllable for all speakers except for the nasal of speaker 1 (56~66 ms). The enclitic stressed syllable of the enclitic structure as well as its vowel were longer than the ultimate, unstressed syllable of the proclitic structure for all speakers whereas its nasal was longer only for speakers 2 (70~66), 3 (64~62) and 4 (54~50 ms).

Postfocal distribution

In postfocal position (Table 2.4, 3e~4e), on the syntagmatic plane, the proclitic structure's word stressed~unstressed syllables showed significant differences for both nasals and vowels. The first member's word stressed syllable of the [ˈmaθima tis ~ ˈmaθiˈma tis] contrastive pair was 26 ms longer than its unstressed syllable (150~124 ms, $t(4)=9.7$, $p<0.0005$), 5 ms longer for the nasal (60~55 ms, $t(4)=2.4$, $p<0.05$) and 21 ms longer for the vowel (90~69 ms, $t(4)=17.6$, $p<0.0005$). The enclitic structure's word stressed ~ enclitic stressed syllables showed no significant differences for either nasals or vowels. The second member's enclitic stressed syllable was 14 ms longer than its word stressed syllable (149~135 ms, $t(4)=1.6$, $p>0.05$), 5 ms longer for the nasal (59~54 ms, $t(4)=1.2$, $p>0.05$) and 9 ms longer for the vowel (90~81 ms, $t(4)=1.6$, $p>0.05$).

On the paradigmatic plane, the proclitic structure's word stressed syllable and its segments were significantly longer than enclitic structure's ones. The proclitic structure's was 15 ms longer than the enclitic structure's word stressed syllable (150~135 ms, $t(4)=3.7$, $p<0.01$), 6 ms longer for the nasal (60~54 ms, $t(4)=3.6$, $p<0.025$) and 9 ms longer for the vowel (90~81 ms, $t(4)=3.1$, $p<0.025$). The enclitic structure's enclitic stressed syllable as well as its vowel were significantly longer than the proclitic's unstressed syllable. The enclitic stressed syllable of the enclitic structure was 25 ms longer than the comparable unstressed syllable of the proclitic structure (149~124 ms, $t(4)=3.8$, $p<0.01$), 4 ms longer for the nasal (59~55 ms, $t(4)=0.8$, $p>0.05$), and 21 ms longer for the vowel (90-69 ms, $t(4)=6.8$, $p<0.005$). The proclitic ~ enclitic's unstressed syllable [θi] was 3 ms longer in favour of the enclitic structure (96~93 ms) and the proclitic ~ enclitic [tis] was 7 ms longer in favour of the enclitic structure as well (195~188 ms).

Both nasal and vowel of the word stressed syllable of the proclitic structure were longer than the nasal and vowel of its ultimate, unstressed syllable, for all speakers except for the nasal of speaker 4 (52 ms). The enclitic stressed syllable of the enclitic structure was longer than its word stressed syllable for speakers 2 (166~162 ms), 3 (158~124 ms) and 4 (162~130 ms), but almost the same for speaker 5 (112~110 ms); its vowel was longer for speakers 1 (92~88), 3 (92~74 ms) and 4 (106~84 ms), but almost the same for speaker 5 (68~66 ms); and its nasal was longer for speakers 2 (74~66 ms), 3 (66~50 ms) and 4 (56~46 ms).

Contextual distribution

On the contextual plane (Figure 2.8), the proclitic structure's word stressed syllable had significant differences. The focal word stressed syllable was 17 ms longer than the prefocal one (177~160 ms, $t(4)=5.1$, $p<0.005$), 4 ms longer for the nasal (66~62 ms, $t(4)=0.9$, $p>0.05$) and 13 ms longer for the vowel (111~98 ms, $t(4)=4.8$, $p<0.005$), and 27 ms longer than the postfocal one (177~150 ms, $t(4)=3.2$, $p<0.025$), 6 ms longer for the nasal (66~60 ms, $t(4)=1.4$, $p>0.05$) and 21 ms longer for the vowel (111~90 ms, $t(4)=4.3$, $p<0.01$); the prefocal word stressed syllable was 10 ms longer than the postfocal one (160~150 ms, $t(4)=1.5$, $p>0.05$), the difference being on the vowel (98~90 ms, $t(4)=1.3$, $p>0.05$).

The enclitic structure's contextual word stressed syllables had no significant differences whereas the contextual comparisons of the enclitic stressed had significant differences between the focal~non focal productions. The focal word stressed syllable was 2 ms longer than the prefocal syllable (137~135 ms) and 2 ms longer than the postfocal one (137~135 ms). The focal enclitic stressed syllable was 16 ms longer than the prefocal enclitic stressed syllable (162~146 ms, $t(3)=1.3$, $p>0.05$), 2 ms longer for the nasal (57~55 ms) and 14 ms longer for the vowel (105~91 ms, $t(3)=1.5$, $p>0.05$), and 13 ms longer than the postfocal one (162~149 ms, $t(3)=2.9$, $p<0.025$), 2 ms shorter for the nasal (57~59 ms) and 15 ms longer for the vowel (105~90 ms, $t(3)=3.4$, $p<0.025$); the prefocal~postfocal enclitic stressed syllables were of about the same duration (146~149 ms).

All speakers produced the word stressed syllable and its vowel in focus with longer duration than out of focus in a constant way. On the other hand, the focal nasal was shorter than the prefocal nasal for speakers 1 (70~73 ms) and 5 (56~58 ms) as well as shorter than the postfocal nasal for speaker 1 (70~74 ms) and equally long for speaker 5 (56 ms).

Most speakers produced the enclitic stressed syllable and its vowel in focus with longer duration than out of focus; the focal syllable was shorter than the prefocal one for speaker 2 (186~202 ms) and almost equally long as the postfocal one for speaker 3 (160~158 ms) whereas the focal vowel was shorter than the prefocal one for speaker 2 (116~120 ms). The nasal was quite variable among speakers, as no speaker produced the focal nasal with longer duration than both the prefocal and postfocal nasals in a constant way.

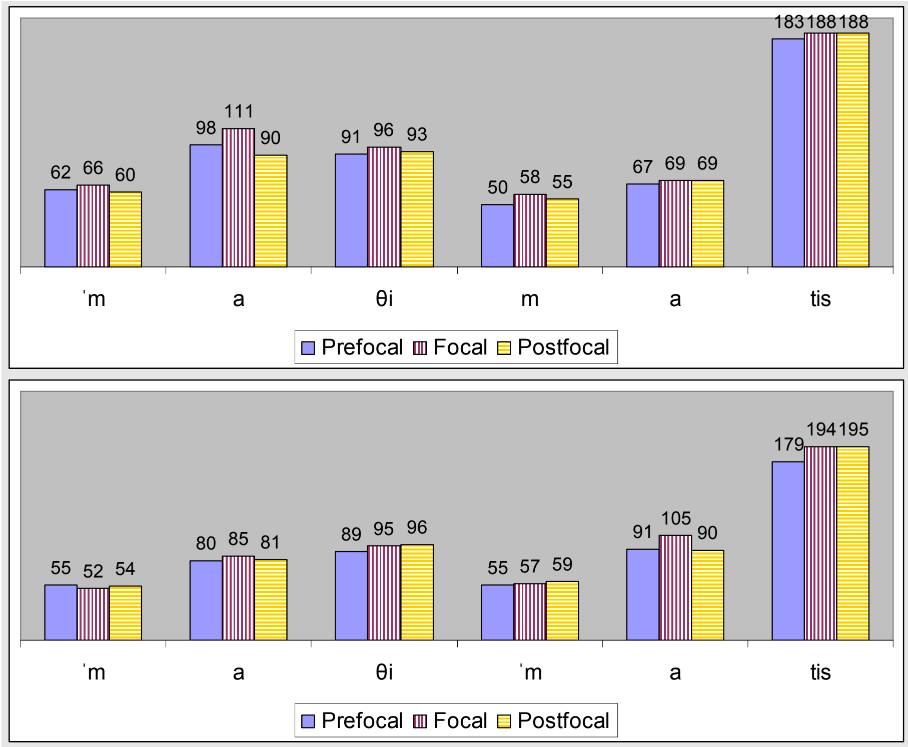


Figure 2.8. Five speakers' average durations (grand mean) in ms of the minimal pair [ˈmaθima tis~ˈmaθiˈma tis] in the carrier sentence [to ˈneo ___ ˈíne ˈðiskolo] in prefocal, focal and postfocal position.

To summarize our results on the duration pattern of word and enclitic stress, the word stressed syllable of the proclitic structure as well as its segments were longer than the unstressed ones whereas the focal distinction was mainly carried out by the vowel. The enclitic stressed syllable of the enclitic structure showed a tendency to be longer than the word stressed syllable but did not reach the significance level (0.05) in this experiment, and this tendency is apparent on the vowels. The enclitic stressed syllable in focus was usually longer than out of focus and this distinction was carried out by the vowel; the durational differences for the consonant were often in the opposite direction. The word stressed syllable of the enclitic structure as well as its segments were shorter than the word stressed syllable of the proclitic structure whereas the enclitic stressed syllable of the enclitic structure was longer than the unstressed, ultimate syllable, of the proclitic structure and this distinction was mainly carried out by the vowel.

Considering the prefocal and postfocal enclitic stressed syllables of the enclitic structure as a whole, with the proclitic's unstressed ultimate as reference, the enclitic stressed syllables are 22% (120~147 ms) longer than the unstressed ones out of focus and 35% longer (120~162 ms) in focus.

2.3.2.2 Fundamental Frequency

Figure 2.9 shows one speaker's raw F0 contour of the proclitic~enclitic structure [ˈmaθima tis~ˈmaθiˈma tis] in their carrier sentences with different contextual frames. It is evident that the utterances are divided into three parts according to their contextual frame: A prefocal part, an F0 rise aligned with word and/or enclitic stress; a focal part, an extended F0 rise followed by an abrupt fall associated with the sentence stress; and a postfocal part, a falling and flattening F0 to the end of the utterance.

Some utterances of the proclitic~enclitic structure, although with different contextual frames, have quite similar F0 structure. Thus, the proclitic's production 3a, with focus on the adjective [ˈðiskolo], has the same F0 structure with 3d which has focus on the proclitic [tis]; production 3b, with focus on the noun [ˈmaθima] and the proclitic [tis] has the same structure as 3c, with focus only on [ˈmaθima]. On the other hand, the enclitic's productions 4b, 4c, 4d appear with quite the same prosodic manifestation. We will concentrate on 3a, 3b, 3e ~ 4a, 4b, 4e leaving out 3c, 3d ~ 4c, 4d, a matter we will take up in the next section (see discussion).

Prefocal distribution

In prefocal position (Table 2.5, 3a~4a), on the syntagmatic plane, the five speakers' mean F0-contour of the first member of the minimal structure [ˈmaθima tis~ˈmaθiˈma tis] was 120 Hz at the beginning of the nasal and 134 Hz at the end of the vowel of the word stressed syllable, a rise of 14 Hz; the F0-contour continued rising up to 148 Hz and remained rather level during the ultimate syllable. In the second member of the pair, the F0-contour rose 8 Hz, from 127 Hz at the beginning of the nasal to 135 Hz at the end of the vowel of the word stress syllable; in the enclitic stress syllable, the F0-contour rose 17 Hz, from 112 Hz at the beginning of the nasal to 129 Hz at the end of the vowel. On the paradigmatic plane, the word stress of the proclitic structure had an F0-rise of 14 Hz whereas the word stress of the enclitic structure had a rise of 8 Hz; on the other hand, the proclitic structure had a level F0-contour on the ultimate unstressed syllable whereas the enclitic structure had a rise of 17 Hz at the ultimate enclitic stress syllable.

All speakers produced the word stressed syllable of the proclitic structure with an F0-rise as a part of an F0 gesture which started at the beginning of the stressed syllable and reached the highest point at the one of the two post-stressed syllables. On the other hand, for speakers 1 and 2 the word stress rise was tonally assimilated to the upcoming focal gesture in a rather rising and flattening F0-contour whereas for speakers 3, 4 and 5 the word stress F0-rise formed part of an independent F0 gesture which ended at the end of the word stress group with an F0-fall.

The word stress group of the enclitic structure was produced as a tonally independent stress group by speakers 2, 3 and 5 whereas speaker 1 assimilated it to the upcoming focal gesture and speaker 4 had an F0-fall within the stress group and was excepted from the other speakers. The F0-maximum of the word stressed syllable was lower than the phrase stressed syllable for speaker 1 (93~99 Hz), higher for speaker 2 (173~140 Hz), and equal for speaker 3 (144~146 Hz) and 5 (130~131 Hz).

Focal distribution

In focal position (Table 2.5, 3b~4b), on the syntagmatic plane, the proclitic structure's word stressed syllable had an extended F0 rise of 53 Hz, from 146 Hz at the beginning of the nasal to 199 Hz at the end of the vowel and fell down to 107 Hz at the beginning of the ultimate to 96 Hz at the end of the ultimate. The enclitic structure's word stressed syllable had an F0-rise of 16 Hz, from 145 Hz at the beginning of the nasal to 161 Hz at the end of the vowel; the enclitic stressed syllable had an extended F0 rise of 44 Hz, from 153 Hz at the beginning of the nasal to 197 Hz at the end of the vowel. On the paradigmatic plane, the word stress of the proclitic structure had a rise of 53 Hz and the word stress of the enclitic structure had a rise of 16 Hz; the unstressed ultimate of the proclitic structure had a fall of 11 Hz as a part of a larger postfocal fall whereas the phrase stressed syllable of the enclitic structure had a rise of 44 Hz.

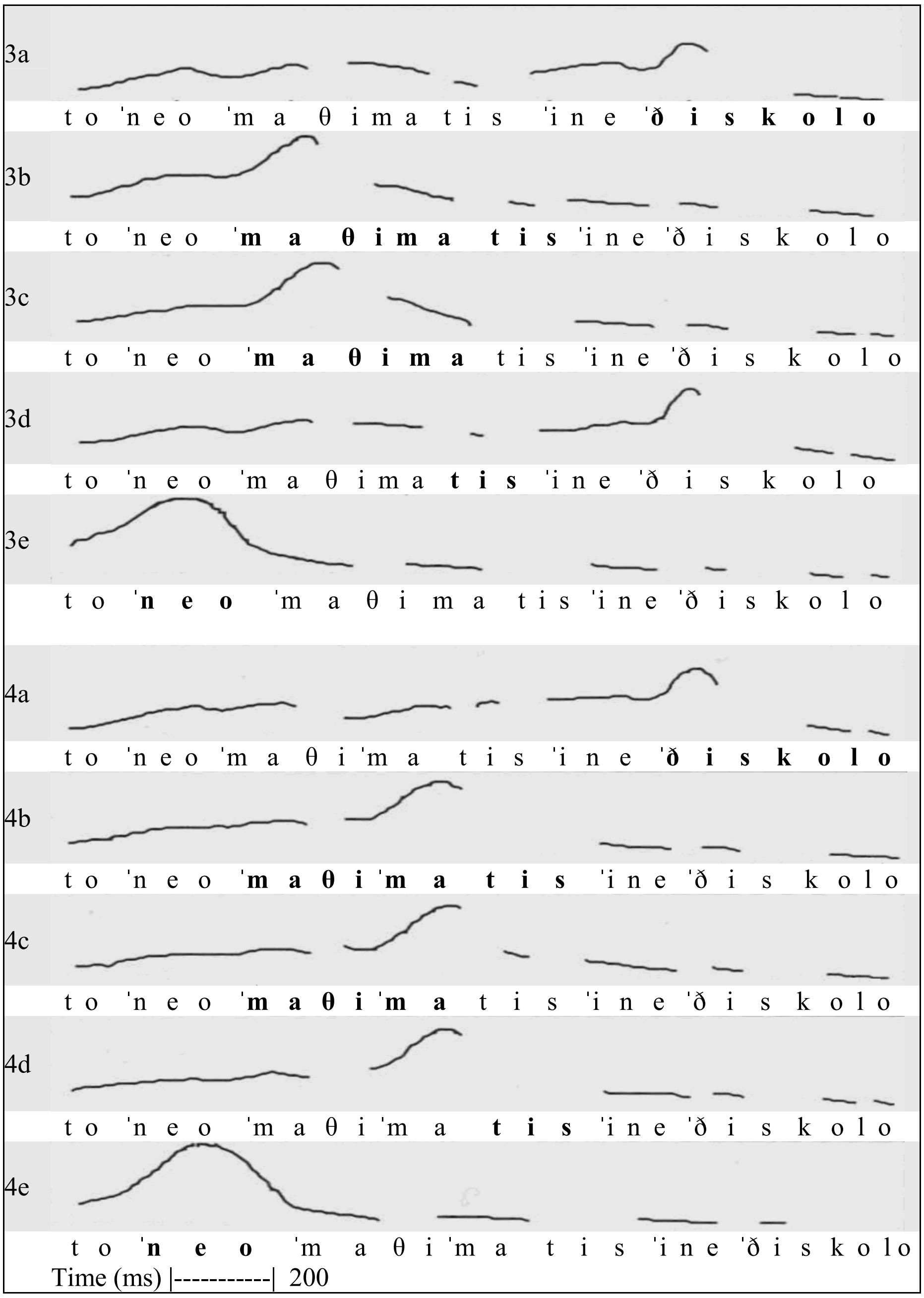


Figure 2.9. One speaker's raw F0 contours of the minimal pair [ˈmaθima tis ~ ˈmaθiˈma tis] (3-4) in the carrier sentence [to ˈneo ___ ˈine ˈðískolo] with different focal organizations (a-e).

Table 2.5, Five speakers' mean voice fundamental frequency (Hz), standard deviation (sd) and grand mean ($\bar{\bar{X}}$G) of the minimal pair [ˈmaθima tis ~ ˈmaθiˈma tis] in the carrier sentence [to ˈneo ___ ˈine ˈðiskolo] in prefocal (3a~4a), focal (3b~4b), and postfocal position (3e~4e). *Speaker 4 is not included in the calculation of the grand mean of the enclitic structure in 3a~4a.

(3a~4a)		ˈm	a	θ i	m	a	ˈm	a	θ i	ˈm	a
1	Hz	100	108		115	119	93	93		94	99
	sd	11.7	12.8		11.7	8.2	4.2	4.2		6.2	12.2
2		124	132		149	180	168	173		114	140
		14.7	10.9		18.8	3.5	7.5	13.0		8.9	14.5
3		88	109		146	140	130	144		118	146
		2.7	4.1		4.1	3.5	3.5	4.1		2.7	4.1
4*		168	194		200	186	(208	164		143	184)
		8.3	8.9		7.0	5.4	(8.3	5.4		8.3	11.4)
5		122	131		130	126	120	130		123	131
		2.7	2.5		3.5	4.1	1.0	1.0		1.3	0.8
	$\bar{\bar{X}}$G*	120	134		148	150	127	135		112	129
	sd	30.6	35.0		32.1	30.9	31.0	33.2		12.7	20.9
(3b~4b)		**ˈm**	**a**	**θ i**	**m**	**a**	**ˈm**	**a**	**θ i**	**ˈm**	**a**
1	Hz	111	237		98	84	91	96		104	215
	sd	16.7	4.4		2.7	4.1	2.1	7.5		9.2	10.0
2		169	201		117	99	166	200		187	226
		3.6	4.1		2.1	2.6	6.5	6.1		6.7	5.4
3		143	184		89	83	153	173		155	181
		4.4	6.5		1.0	1.0	2.7	5.7		5.0	5.4
4		190	232		133	120	193	205		190	216
		7.0	8.3		2.7	5.3	4.4	10.0		12.2	18.1
5		121	142		100	94	124	134		131	147
		2.0	2.5		1.4	0.5	1.0	3.7		2.1	2.5
	$\bar{\bar{X}}$G	146	199		107	96	145	161		153	197
	sd	32.8	38.7		17.5	15.0	39.2	46.2		36.7	32.7
(3e~4e)		ˈm	a	θ i	m	a	ˈm	a	θ i	ˈm	a
1	Hz	98	90		87	83	98	89		84	81
	sd	4.1	0.8		2.1	2.7	2.0	2.6		1.3	1.6
2		119	100		95	90	116	101		100	98
		2.9	1.4		1.7	1.0	2.0	2.0		1.6	1.4
3		96	83		83	81	95	80		87	80
		2.9	2.2		2.6	1.6	1.7	1.0		2.7	1.4
4		130	118		114	109	127	116		114	110
		3.4	3.4		2.9	1.7	3.7	3.5		3.2	2.9
5		110	100		95	90	107	95		91	89
		3.9	3.8		2.1	1.4	2.1	2.5		2.0	1.6
	$\bar{\bar{X}}$G	110	98		94	90	108	96		95	91
	sd	14.3	13.1		11.9	11.0	13.1	13.5		12.1	12.5

All speakers produced the word stressed syllable of the proclitic structure in focus with a large F0 rise and with the highest F0 top as well as with the largest F0 fall into a low F0 level in the post-stress syllable in a constant way. The enclitic stressed syllable in focus appeared with the highest F0 top within the utterances as well as with a larger F0 rise in comparison to the word stressed syllable of the enclitic structure. Moreover, the enclitic structure appeared with two tonal groups for all speakers but speaker 1 whose word stress group was tonally assimilated to the enclitic stress group; speakers 2 and 5 showed an anathetic F0 contour whereas speakers 3 and 4 an alternating F0-contour.

Postfocal distribution.

In postfocal position (Table 2.5, 3e~4e), the F0 contour of the proclitic ~ enclitic structure was quite the same, with no considerable variation between the stressed ~ unstressed or word stressed ~ enclitic stressed syllables. It declined from 110 to 98 Hz during the word stressed syllable and from 94 to 90 Hz during the ultimate of the proclitic structure, and from 108 to 96 Hz during the word stressed syllable and from 95 to 91 Hz during the enclitic stressed syllable of the enclitic structure.

All speakers produced the proclitic ~ enclitic structure with a declining and flattening F0 contour at a low level except for speaker 3 who showed an intersyllabic level F0 contour (83 Hz) for the proclitic structure as well as an intersyllabic resetting (95-80~87-80 Hz) for the enclitic structure.

Contextual distribution

On the contextual plane (Figure 2.10), the proclitic structure's word stressed syllable in focus started 26 Hz higher than the same syllable in prefocal position (146-120 Hz) and 36 Hz higher than the postfocal one (146~110 Hz), rose 65 Hz higher than the prefocal word stressed syllable (199~134 Hz) and 101 Hz higher than the postfocal one (199~98 Hz); the ultimate unstressed syllable was at a high F0 level in prefocal position (148~150 Hz), whereas at a low level in focal (107~96 Hz) and postfocal position (94~90 Hz). The enclitic's word stressed syllable in focus started 18 Hz higher than the prefocal (145~127 Hz) and 37 Hz higher than the postfocal one (145~108 Hz), rose 26 Hz higher than the prefocal (161~135 Hz) and 65 Hz higher than the postfocal one (161~96 Hz); the enclitic stressed syllable in focus started 41 Hz higher than the prefocal (153~112 Hz) and 58 Hz higher

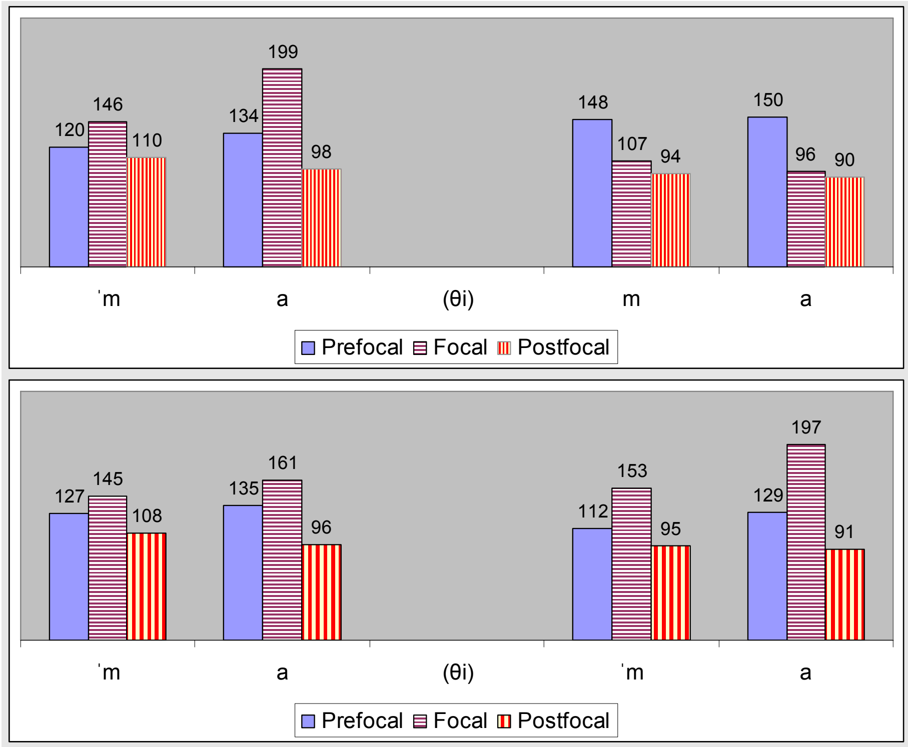


Figure 2.10. Five speakers' average voice fundamental frequency (grand mean) in Hz of the minimal pair [ˈmaθima tis ~ ˈmaθiˈma tis] in the carrier sentence [to ˈneo ___ ˈine ˈðiskolo] in prefocal, focal and postfocal position.

than the post focal one (153~95 Hz), and rose 68 Hz higher than the prefocal (197~129 Hz) and 106 Hz higher than the postfocal one (197~91 Hz).

All speakers produced the word stressed syllable of the proclitic structure as well as the enclitic stressed syllable of the enclitic structure in focus with a larger F0 rise, a higher F0 top and a larger F0 fall than in prefocal position except for speaker 3 who produced the prefocal enclitic stressed syllable with an F0 rise about equal to that of the focal enclitic stressed syllable (118-146~155-181 Hz).

The word stress F0 rise of the proclitic structure in focus started at a higher level than the prefocal one except for speaker 5 who started the F0 rise at the same level (121~122 Hz). On the other hand, when the enclitic structure was in focus, the word stress F0 rise started at about the same level for speakers 1, 2, 5 and at 23 Hz higher than the prefocal word stress rise for speaker 3 (153-130 Hz) whereas the enclitic stress F0 rise started higher than the prefocal rise for all speakers.

To summarize the results of the F0 pattern of word stress and enclitic stress, both appeared either as independent tonal groups associated with the corresponding stress groups or assimilated to the upcoming tonal (focal) group. Both word stressed and enclitic stressed syllables in focus appeared with the largest F0 rise, the highest F0 top and the largest F0 fall within the utterance as well as with a somewhat raised F0 contour. Moreover, the word stress tonal group appeared in an alternating F0 contour, either anathetic, or flattening, in accordance with its degree of assimilation to the focal stress group whereas the word stress group had no effect on a larger domain. Finally, both postfocal word stress and enclitic stress appeared with the minimal F0 variation at the lowest level within the utterance.

2.3.2.3 Intensity

Figure 2.11 shows one speaker's raw intensity contours of the proclitic~enclitic structure [ˈmaθima tis ~ ˈmaθiˈma tis] in their carrier sentences in different contextual frames. The proclitic structure appears with the typical word stress prosodic realization, i.e. stressed syllables have higher intensity than unstressed ones, both in focus (3b, 3c, 3d) and out of focus (3a, 3e). The enclitic structure for this speaker has equal peak intensity contours for word stress and enclitic stress in prefocal position (4a), somewhat higher for the enclitic stress in focal position (4b, 4c, 4d), and lower than the word stress in postfocal position (4e); the intensity contour is also lowered in postfocal position.

Prefocal distribution

In prefocal position (Table 2.6, 3a~4a), on the syntagmatic plane, the word stressed syllable of the proclitic structure was -6.7 dB and its unstressed ultimate -10.8 dB, a significant difference of 4.1 dB ($t(4)=5.3$, $p<0.005$). The word stress syllable of the enclitic structure was -9.3 dB and its enclitic stress syllable was -11.1 dB. On the paradigmatic plane, the proclitic~enclitic word stress difference was 2.6 dB (-6.7~-9.3, $t(3)=0.5$, $p>0.05$) whereas the proclitic's ultimate had about the same intensity as the enclitic's enclitic stressed syllable (-10.8~-11.1 dB).

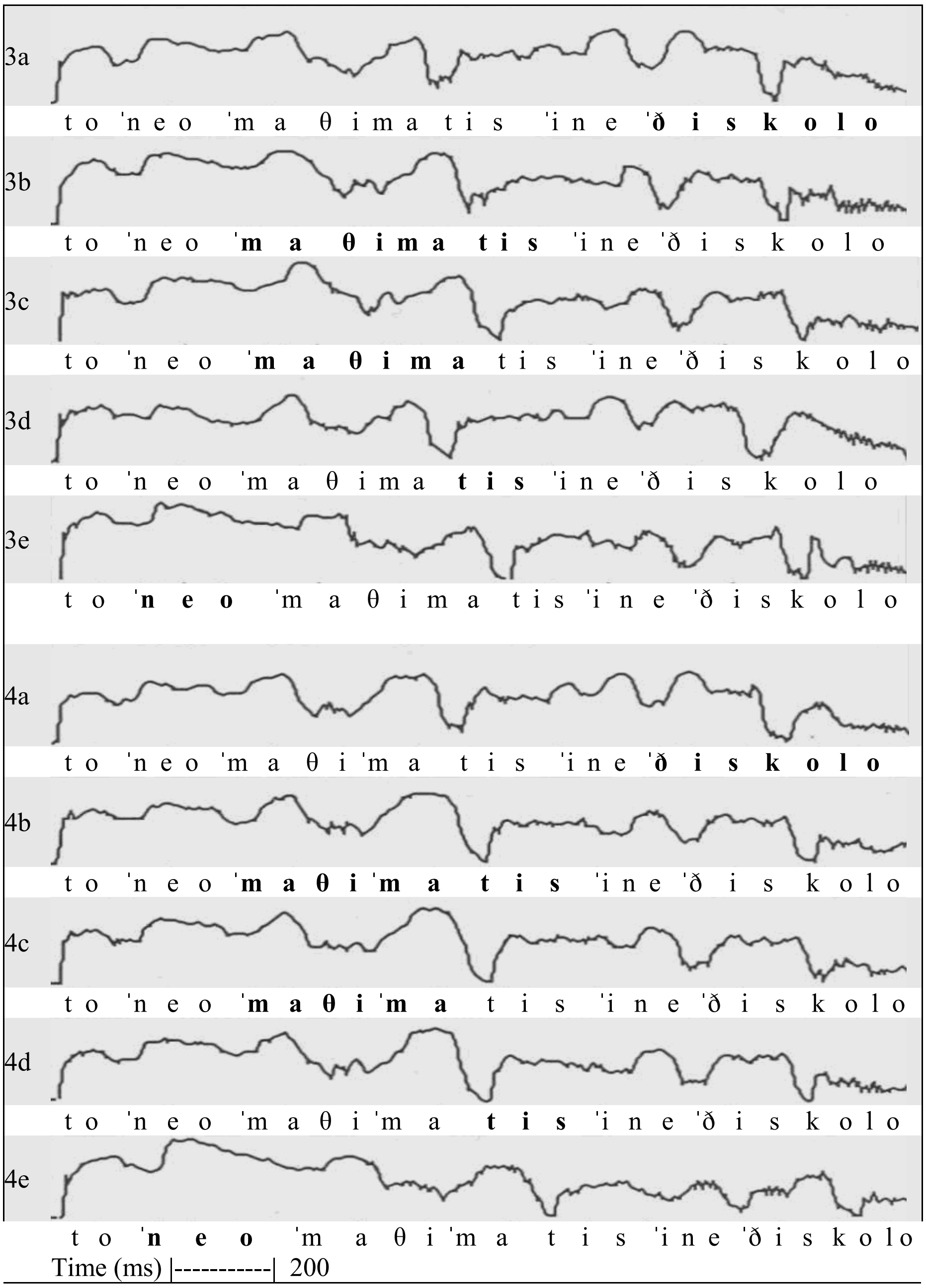


Figure 2.11. One speaker's intensity contours of the minimal pair [ˈmaθima tis ~ ˈmaθiˈma tis] (3~4) in the carrier sentence [to ˈneo ___ ˈine ˈðiskolo] with different focal organizations (a-e).

Table 2.6. Five speakers' mean intensity (dB), standard deviation (sd) and grand mean ($\bar{\bar{X}}$G) of the minimal pair [ˈmaθima tis ~ ˈmaθiˈma tis] in the carrier sentence [to ˈneo ___ ˈine ˈðiskolo] in prefocal (3a~4a), focal (3b~4b), and postfocal position (3e~4e). *Speaker 4 is not included in the calculation of the grand mean of the enclitic structure in 3a~4a.

(3a~4a)		ˈma	θi	ma	ˈma	θ i	ˈma
1	dB	-6.6		-12.8	-11.6		-14.6
	sd	3.3		2.8	4.4		4.2
2		-9.0		-10.8	-4.2		-8.6
		1.4		1.0	1.7		2.6
3		-11.6		-16.4	-13.2		-12.6
		2.7		2.0	0.8		0.8
4*		-1.0		-6.0	(-2.2		-2.2)
		0.0		1.0	(0.8		0.8)
5		-5.4		-8.4	-8.2		-8.8
		0.5		0.5	0.4		0.4
	* $\bar{\bar{X}}$G	-6.7		-10.8	-9.3		-11.1
		3.9		4.0	3.9		2.9
(3b~4b)		**ˈma**	**θi**	**ma**	**ˈma**	**θi**	**ˈma**
1	dB	-1.0		-17.4	-12.2		-1.0
	sd	0.0		1.9	4.2		0.0
2		-1.0		-15.6	-6.0		-5.8
		0.0		1.5	1.2		0.8
3		-5.4		-19.2	-9.0		-9.8
		0.5		1.0	2.2		2.0
4		-1.0		-12.6	-3.4		-1.0
		0.0		0.8	1.6		0.0
5		-4.2		-15.0	-5.4		-5.4
		1.0		1.4	0.5		0.5
	$\bar{\bar{X}}$G	-2.5		-15.9	-7.2		-4.6
	sd	2.1		2.4	3.4		3.7
(3e~4e)		ˈma	θi	ma	ˈma	Θi	ˈma
1	dB	-6.4		-21.8	-11.0		-21.2
	sd	1.8		2.0	3.0		1.0
2		-9.8		-15.4	-10.4		-12.8
		1.4		0.8	0.8		0.8
3		-13.6		-20.4	-16.6		-17.2
		1.3		2.3	0.5		0.8
4		-3.8		-12.4	-5.4		-6.4
		1.0		1.6	1.1		1.1
5		-6.6		-15.6	-13.2		-13.2
		1.3		1.6	2.0		1.7
	$\bar{\bar{X}}$G	-8.0		-17.1	-11.3		-14.1
	sd	3.7		3.8	4.1		5.5

All speakers produced the word stress syllable of the proclitic structure with higher intensity than its unstressed ultimate with high constancy. The word and enclitic stress syllables of the enclitic structure showed some variation, speakers 1 and 2 had higher intensity for the word stress syllable (Sp. 1 -11.6~-14.6, Sp. 2 -4.2~-8.6 dB), speakers 3 and 5 about equal intensity for both the word stress and enclitic stress syllables (Sp. 3 -13.2~-12.6, Sp. 5 -8.2~-8.8 dB). The intensity of the word stress syllable of the proclitic structure was higher than the word stress syllable of the enclitic structure except for speaker 2 (-9.0~-4.2 dB) whereas its unstressed ultimate was higher for speaker 1 (-12.8~-14.6 dB), lower for speakers 2 and 3 (Sp. 2 -10.8~-8.6, Sp. 3 -16.4~-12.6 dB) and almost equal for speaker 5 (-8.4~-8.8 dB).

Focal distribution

In focal position (Table 2.6, 3b~4b), on the syntagmatic plane, the proclitic's word stress syllable was -2.5 dB and its unstressed ultimate -15.9 dB, a significant difference of 13.4 dB ($t(4)=13.2$, $p<0.0005$); the enclitic's word stress syllable was -7.2 dB and its enclitic stress syllable -4.6 dB ($t(4)=-1.1$, $p>0.05$). On the paradigmatic plane, the proclitic ~ enclitic word stress syllables had a significant difference of 4.7 dB (-2.5~-7.2, $t(4)=2.6$, $p<0.05$) and the ultimate unstressed ~ enclitic stress syllables, a significant difference of 11.3 dB (-15.9~-4.6, $t(4)=-8.6$, $p<0.0005$).

All speakers produced the word stress syllable of the proclitic structure with higher intensity than its unstressed ultimate. The enclitic stress syllable of the enclitic structure had higher intensity for speaker 1 (-1.0~-12.2 dB) and speaker 4 (-1.0~-3.4 dB) and almost equal intensity with its word stress syllable for speaker 2 (-5.8~-6.0 dB), speaker 3 (-9.8~-9.0 dB) and speaker 5 (-5.4~-5.4 dB). On the other hand, all speakers had higher intensity for the word stress syllable of the proclitic structure than for the word stress syllable of the enclitic structure as well as having higher intensity for the enclitic stress syllable of the enclitic structure than for the unstressed ultimate of the proclitic structure.

Postfocal distribution

In postfocal position (Table 2.6, 3e~4e), on the synagmatic plane, the proclitic's stressed ~ unstressed syllables had a significant difference of 9.1 dB (-8.0~-17.1, $t(4)=5.3$, $p<0.005$); the enclitic's word stress~enclitic stress syllable had a difference of 2.8 dB (-11.3~-14.1, t(4)=1.5, p>0.05). On the paradigmatic plane, the proclitic’s word stress syllable was 3.3 dB higher then the enclitic’s word stress syllable (-8.0~-11.3, t(4)=3.0, p<0.025) whereas the proclitic’s ultimate was 3 dB lower than the enclitic’s enclitic stress syllable (-17.1~-14.1, t(4)=-3.3, p<0.025).

All speakers produced the word stress syllable of the proclitic structure with higher intensity than its unstressed ultimate with high constancy. The word and enclitic stress syllables of the enclitic structure showed some variation, speakers 1 and 2 had higher intensity for the word stress syllable (Sp. 1 -11.0~-21.2, Sp. 2 -10.4~-12.8 dB) and speakers 3,4 and 5 about equal intensity for the word stress and enclitic stress syllables (Sp. 3 -16.6~-17.2, Sp. 4 -5.4~-6.4, Sp. 5 -13.2~-13.2 dB). The intensity of the word stress syllable of the proclitic structure had higher intensity than the word stress syllable of the enclitic structure for all speakers but speaker 2 (-9.8~-10.4 dB) whereas its unstressed ultimate had lower intensity than the enclitic stress syllable of the enclitic structure for all speakers but speaker 1 (-21.8~-21.2 dB).

Contextual distribution

On the contextual plane (Figure 2.12), the proclitic’s word stress syllable of the prefocal~focal opposition showed a significant difference of 4.2 dB (-6.7~-2.5, t(4)=-2.7, p<0.05), the focal~postfocal opposition 5.5 dB (-2.5~-8.0, t(4)=4.1, p<0.01), and the prefocal~postfocal opposition a difference of 1.3 dB (-6.7~-8.0, t(4)=2.5, p<0.05); the unstressed ultimate of the prefocal ~ focal opposition showed a significant difference of 5.1 dB (-10.8~-15.9, t(4)=7.1, p<0.005), the focal~postfocal 1.2 dB (-15.9~-17.1, t(4)=1.3, p>0.05), and the prefocal~postfocal 6.3 dB (-10.8~-17.1, t(4)=6.9, p<0.005). The enclitic’s word stress syllable of the prefocal ~ focal opposition showed a difference of 2.1 dB (-9.3~-7.2, t(3)=-0.8, p>0.05), the focal~postfocal 4.1 dB (-7.2~-11.3, t(4)=2.4, p<0.05), and the prefocal~ postfocal 2 dB (-9.3~-11.3, t(3)=2.3, p<0.05); the enclitic stress syllable of the prefocal~focal opposition showed a variation of 6.5 dB (-11.1~-4.6, t(3)=-2.1, p>0.05), the focal~postfocal 9.5 dB (-4.6~-14.1, t(4)=3.5, p<0.025), and the prefocal~postfocal 3 dB (-11.1~-14.1, t(3)=8.9, p<0.005).

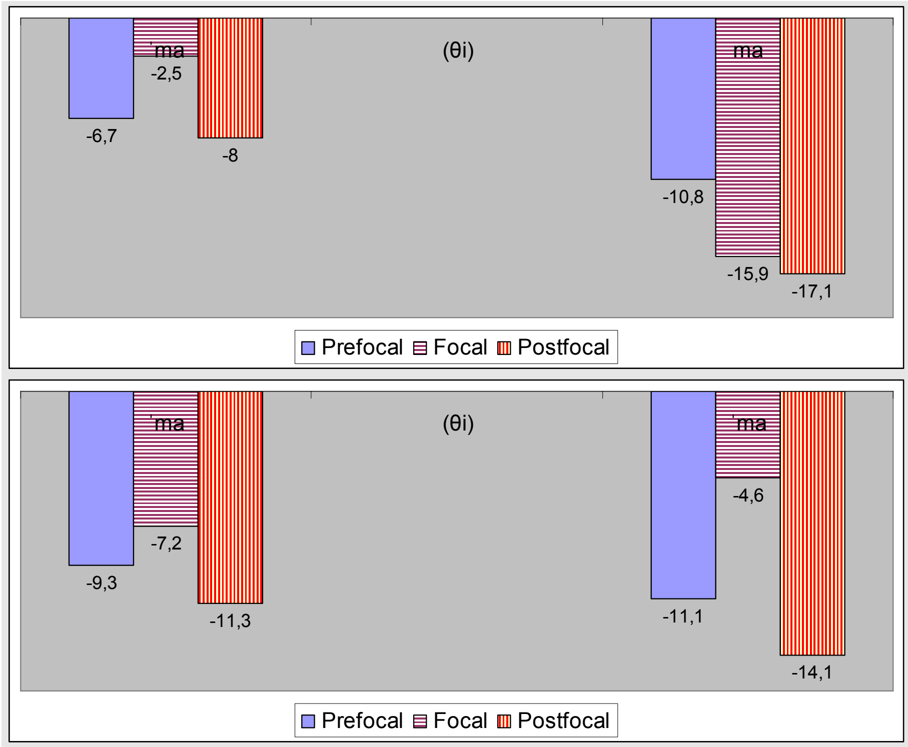


Figure 2.12. Five speakers' average intensity (grand mean) in dB of the minimal pair [ˈmaθima tis ~ ˈmaθiˈma tis] in the carrier sentence [to ˈneo ___ ˈine ˈðiskolo] in prefocal, focal and postfocal position.

All speakers produced the word stressed syllable of the proclitic structure in focus with higher intensity than out of focus. The difference between the word stress~unstressed ultimate syllable in focus was also larger than out of focus for all speakers but speaker 1 who produced the focal~postfocal opposition with almost equal intensity (-16.4~-15.4 dB). All speakers produced the enclitic stressed syllable of the enclitic structure in focus with higher intensity than out of focus as well, although the prefocal~focal opposition did not reach the significance level (0.05). On the other hand, the difference between the word stressed ~ enclitic stressed syllables showed some variation among the speakers in the three different focal positions, the largest one for speaker 1 who produced the focal enclitic stressed syllable with 11.2 dB higher intensity than the word stress syllable (-1.0~-12.2 dB), the prefocal one with 3 dB lower intensity (-14.6~-11.6 dB), and the postfocal one with 10.2 dB lower intensity (-21.2~-11.0 dB).

To summarize the results of the intensity pattern of word stress and enclitic stress, the word stressed syllable of the proclitic structure was produced with higher intensity than its unstressed ultimate and this intensity difference was even larger in focal position. The intensity of the enclitic stressed syllable of the enclitic structure was higher than the intensity of the word stressed syllable only in focal position whereas the intensity ratio between the word and enclitic stress was smaller than the intensity ratio of the contrastive proclitic structure in non-focal positions.

2.3.3 Discussion

To recapitulate the results of this experiment, enclitic stress is realized by the synergetic contribution of duration, F0 and intensity before focus, duration and intensity after focus. In focal position, the enclitic stress neutralizes the application of sentence stress to any other element of the enclitic structure, realizing itself as an extended increase in duration, F0, and intensity locally and as a partial reorganization of the prosodic manifestation (mainly F0) globally. The proclitic ~ enclitic opposition lies mainly in the presence or absence of the enclitic stress associated with the enclitic structure, rather than the proclitic or enclitic elements as such[6].

In prefocal position, the enclitic stressed syllables have not been found significantly longer than the word stressed syllables whereas the word stressed syllables of the enclitic structure are shorter than the word stressed syllables of the proclitic structure, i.e. we have a partial weakening of the word stress in an enclitic structure, which is mainly kept on the paradigmatic plane. The enclitic's ultimate, enclitic stressed syllable, is significantly longer than the proclitic's ultimate, unstressed syllable, i.e. the enclitic stress duration overrides any word and phrase final lengthening. Moreover, the enclitic's pronoun appears with the same duration as the proclitic's pronoun,

[6]In a listening test, subjects could not differentiate between the proclitic ~ enclitic structures of the type (i ˈθia) (mu ˈpulise to ˈspiti) (the aunt sold me the house) and (i ˈθia mu) (ˈpulise to ˈspiti) (my aunt sold the house), produced at a normal tempo, an indication that enclitic stress is closer to the rhythmic structure of Greek than to syntax; in the latter case prosody may be used to disambiguate sentences with different constituent structures (Lehiste 1973).

contrary to the expectation of the cumulative phrase-final lengthening principle (Paccia-Cooper and Cooper 1981), according to which final syllables at word and phrase level should be lengthened in a rather cumulative manner.

F0 starts rising at the beginning of the enclitic stressed syllable, reaches the highest point at the end of it or on the first post-stress syllable and then falls gradually to the end of the stress group or remains rather level, probably in anticipation of the sentence stress, resembling thus the realization of the word stress, at least in structure. The proclitic ~ enclitic opposition appears as a rather local phenomenon leaving the prosodic context unaffected and eliminating the corresponding prosodic realization within the proper domain, the proclitic structure with one stress group having a word stress and the enclitic structure with two stress groups having one word stress and one enclitic stress.

The top F0 for the word stress and enclitic stress of the enclitic structure does not show any regular declination effect ('t Hart and Cohen 1973, Pierrehumbert 1979, Cooper and Sorensen 1981) between the two prosodic categories. Our Greek data support the view that there is hardly any noticeable F0 declination in prefocal position in Swedish (Bruce 1982), there may even be an F0 inclination up to focus (Gårding et al. 1982), in opposition to Cooper et al.'s (1985) findings that declination is a rule across the whole utterance in English, an indication that the declination effect may vary across utterances as well as across different languages.

The peak intensity of the enclitic stressed syllables is quite close to the intensity of the word stress syllables of the enclitic structure but lower in comparison to the word stressed syllables of the proclitic structure, i.e. apart from duration we have an intensity weakening of the word stress (Malikouti-Drachman 1976) in the enclitic structure, too, an indication that enclitic stress may have a larger domain than the stress group it belongs to, namely the whole phrase.

In focal position, the enclitic stressed syllables are longer whereas the word stressed syllables seem unaffected by the application of the sentence stress in the enclitic structure. The contribution of duration to sentence stress does not appear totally constant across the subjects, an indication that duration is not an absolute acoustic correlate for focus even when it is realized on enclitic stress. The proclitic ~ enclitic elements, even in focus, do not show any major duration difference as a result of the final lengthening reported for English (Klatt 1976). Instead, in the enclitic case, the focal contribution of duration is realized on the enclitic stress, the host of the enclitic element.

The F0 manifestation for focus in an enclitic structure with an enclitic stress has the same basic characteristics that the word stress has, i.e. an extended F0 rise followed by an abrupt F0 fall, an optional F0 compression for the prefocal word stress(es), although an alternating or anathetic F0 contour seems equally preferable across the subjects, and an F0 falling and flattening for any postfocal word stress(es).

Within the enclitic structure, no matter whether the whole structure or parts of it are in focus, sentence stress is placed on the enclitic stress, i.e. a syntactic constraint is imposed on prosody (Rossi 1985). The distinctive load of word stress is rather limited after the application of enclitic stress since for the distribution of the enclitic stress, the position of word stress is an absolute requirement. On the other hand, enclitic stress is highly informative since the morpholexical as well as the syntactic components of the language are involved in its distribution (see part 1).

Comparing the prosodic realization of the proclitic ~ enclitic structures we see different prosodic contributions depending on syntax whereas each syntactic structure may appear with different prosodic realizations depending on the focus ~ presupposition organization of the utterances, i.e. both the syntactic and the infonnative structures interact on prosody (Rossi 1985). The surface prosodic realization of the Greek utterances may well be organized in stress groups, in accordance with Thorsen's (1983) description of Danish. However, the peculiarity of the Greek proclitic ~ enclitic structures has shown that when the enclitic element is in focus, sentence stress is distributed on both the stress group the enclitic belongs to and its syntactic host whereas when the proclitic element is in focus sentence stress is not distributed within the stress group the proclitic belongs to but rather on its corresponding host, another major syntactic constraint of the distribution of sentence stress on prosody.

The different contextual realizations of F0 within the proclitic ~ enclitic structures do not affect the F0 contour either in structure or in F0 range apart from the proclitic element which is realized as a focal F0 change on the host it belongs to, as already mentioned. Our Greek data are in disagreement with Eady et al.'s (1986) who report a different F0 manifestation for narrow ~ broad focus in English.

The intensity of the enclitic stressed syllables in focus is even higher than out of focus. Thus, intensity seems to contribute synergetically along with duration and F0 to the production of sentence stress, realized on enclitic stress, as is the case when sentence stress is realized on word stress. On acoustic grounds it is questionable whether there is enough evidence to differentiate between word and enclitic stress although they have quite different perceptual dimensions (see part 4).

In postfocal position, the enclitic stressed syllables may not be significantly longer than the word stressed syllables whereas the word stressed syllables in the enclitic structure are shorter than the word stressed syllables in the proclitic structure, i.e. the postfocal duration structure of the proclitic ~ enclitic opposition is quite close to the prefocal one. Having found no considerable duration differences for word stress between the prefocal and postfocal positions in the previous experiment either, we may say that the prosodic reorganization (Rossi 1985) is minimal across the Greek utterances, as far as duration is concerned.

The enclitic's enclitic stressed syllable is 22% longer than the proclitic's unstressed ultimate out of focus and 35% longer when in focus. Thus, duration does not seem to contribute more to the enclitic stress than to the word stress since the corresponding lengthening of word stress is 24% out of focus and 39% in focus. The F0 of the postfocal proclitic ~ enclitic structures declines at a low level regardless of word or enclitic stress distinctions. Thus, enclitic stress has a local acoustic realization, quite similar to that of word stress, without having any effect on the rest of the sentence the way sentence stress has.

Öhman et al. (1979) and Lyberg (1981) propose an F0 dependent model for segment duration in Swedish. We have already pointed out the basic independence of the acoustic parameters in the production of word and sentence stress in Greek (Botinis 1982). In the present study, there is further evidence for the prosodic manifestation of enclitic stress, besides word and sentence stress. First, in prefocal position, both word stressed and enclitic stressed syllables appear with rather longer duration whereas the F0 contour may be quite compressed. Second, sentence stress is realized with a large F0 change as a rule whereas duration is not present in a constant way and, at any rate, not to the degree an F0 dependent model should predict. Third, in postfocal position, both word stress and enclitic stress are realized as a rather flat F0 contour whereas duration seems to contribute to these two prosodic categories as a rule. Our Greek data corroborates Bruce's (1981) and Bannert's (1982) proposals for an independent but interrelated model of Swedish and German respectively.

The intensity contour of the postfocal enclitic structure is rather declining in spite of the presence of an enclitic stress on the right. This may be the declination effect which is evident after focus (Bruce 1982), according to which acoustically unequal syllables may be perceptually equally prominent and vice versa (Pierrehumbert 1979), the perceptual system compensating for the declination effect. On the other hand, the proclitic structure has a steeper declination, the absence of an intensity contribution except for the contribution of the word stress. Thus, intensity combined with duration, what has been referred to as the "energy integral", appears to be a solid acoustic parameter for enclitic stress in this position as is the case for word stress (Fischer-Jørgensen 1984).

2.4 Relation between Prosody and Syntax

2.4.0 Introduction

In the third acoustic experiment, the relation between prosody and syntax is investigated. There exist three current views on this topic. First, prosody is dependant on syntax reflecting the sentence's immediate constituent structure. Second, prosody is rather independent from syntax with its own distribution rules and boundaries which, as a rule, do not coincide with the phrase boundaries. Third, there is no one to one correspondence between prosody and syntax but a rather complicated relation adjusted by the association rules. Our main question in this experiment is whether prosody is independent, and if so to what degree, or if it is just a help tool for the language to distinguish speech units with different syntactic and semantic representations.

At this point, we assume that prosody is the product of two basic independent but interrelated components, rhythm and intonation. Primarily, rhythm reflects the morpholexical and syntactic structure of the language and intonation reflects the language's informative and discourse structure. Rhythm divides the utterance through a combination of stressed and unstressed syllables into isochronous stress groups or isochronous syllable groups, i.e. what has been referred to as stress timed vs. syllable timed languages (Pike 1945, Lehiste 1977, Dauer 1983). Similar concepts to stress group (Thorsen 1982) are the foot (Abercrobie 1964) and the groupe prosodique (Di Cristo 1981).

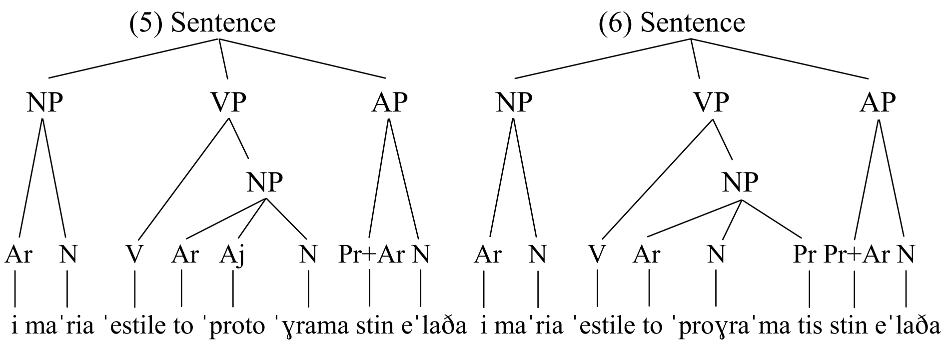
On the other hand, intonation divides the utterance into more or less prominent parts, i.e. a prefocal, a focal and a postfocal part (Bruce 1977), (pre)head, nucleus, tail (Crystal 1969) and attaque, prétonique, tonique (Di Cristo 1981). The isomorphism between stress groups and tonal groups is unequal as has been shown in the two previous acoustic experiments where we found postfocal tonal flattening as well as prefocal tonal assimilations, whereas the focal application to a lexical element within the enclitic structure with an enclitic stress was blocked.

This time only F0 has been considered for the following reasons: First, the acoustic parameters of word and enclitic stress involved in the paratactic ~ enclitic opposition have already been examined in the two previous experiments; second, it has been shown that it is mainly F0 which is most affected by the reorganization of the utterances in accordance with their contextual demands, duration and intensity distribution being quite constant in structure; third, it has been quite difficult to set up natural paratactic ~ enclitic structures with the same segments, easily produced, with maximal, clear-cut semantic oppositions.

2.4.1 Speech Material

Two meaningful Greek sentences were set up; the first sentence [i maˈria ˈestile to ˈproto ˈɣrama stin eˈlaða] (Maria sent the first letter to Greece), and the second sentence [i maˈria ˈestile to ˈproɣraˈma tis stin eˈlaða] (Maria sent her program to Greece). Both sentences were declarative composed of a noun phrase, a verb phrase, and an adverbial phrase. Yet, the noun phrases under the verb phrases were different in syntactic structure; in the first sentence the noun phrase was composed of a definite article, an adjective and a noun whereas, in the second sentence, the noun phrase was composed of a definite article, a noun and a pronoun.

Both sentences had the same number of stresses with, apparently, the same prosodic structure; however, the paratactic structure [to ˈproto ˈɣrama] carries two word stresses into two rather independent lexical elements whereas the enclitic structure [to ˈproɣraˈma tis] carries a word stress and an enclitic stress as a result of the enclitic structure. The crucial question here is if the prosodic categories apply the same way to sentences with different syntactic structures, i.e. if prosody is independent of syntax. The syntactic structures of the two test sentences are shown in (5) and (6).



Each of the test sentences was the carrier of five contextual frames listed in Table II.III. Each test sentence was elicited as an answer formulated in a different way, to make the speaker choose one of the elements of the sentence as the focus and the carrier of the information required by the questioner. Thus, in productions 5a~6a the adverbial phrase [stin eˈlaða] was in focus having the paratactic~enclitic structure [to ˈproto ˈɣrama ~ to ˈproɣraˈma tis] in prefocal position. In 5b~6b the whole noun phrases [to ˈproto ˈɣrama ~ to ˈproɣraˈma tis] were in focus, in 5c~6c the adjective [ˈproto] and the noun [ˈproɣrama] were in focus, and in 5d~6d the noun [ˈɣrama] and the enclitic pronoun [tis] were in focus. Last, in 5e~6e the noun [maˈria] was in focus, having the paratactic~enclitic structure in postfocal position.

Table II.III: The minimal pair [ˈproto ˈɣrama~ˈproɣraˈma tis] in the carrier sentence [i maˈria ˈestile to ___ stin eˈlaða] with different contextual organizations (see text).

Contextual frame	Test sentences
5a.	
[pu ˈestile to ˈproto ˈɣrama i maˈria]	[i maˈria ˈestile to ˈproto ˈɣrama **stin e**ˈ**laða**]
(Where did Maria send the first letter?)	(Maria sent the first letter **to Greece**.)
5b.	
[ti ˈestile i maˈria stin eˈlaða]	[i maˈria ˈestile to ˈ**proto** ˈ**ɣrama** stin eˈlaða]
(What did Maria send to Greece?)	(Maria sent the **first letter** to Greece.)
5c.	
[pço ˈɣrama ˈestile i maˈria stin eˈlaða]	[i maˈria ˈestile to ˈ**proto** ˈɣrama stin eˈlaða]
(Which letter did Maria send to Greece?)	(Maria sent the **first** letter to Greece.)
5d.	
[pço ˈproto ˈestile i maˈria stin eˈlaða]	[i maˈria ˈestile to ˈproto ˈ**ɣrama** stin eˈlaða]
(Which first did Maria send to Greece?)	(Maria sent the first **letter** to Greece.)
5e	
[pça ˈestile to ˈproto ˈɣrama stin eˈlάða]	[i **ma**ˈ**ria** ˈestile to ˈproto ˈɣrama stin eˈlaða]
(Who sent the first letter to Greece?)	(**Maria** sent the first letter to Greece.)
6a.	
[pu ˈestile i maˈria to ˈproɣraˈma tis]	[i maˈria ˈestile to ˈproɣraˈma tis **stin e**ˈ**laða**]
(Where did Maria send her program?)	(Maria sent her program to **Greece**.)
6b.	
[ti ˈestile i maˈria stin eˈlaða]	[i maˈria ˈestile **to** ˈ**proɣra**ˈ**ma tis** stin eˈlaða]
(What did Maria send to Greece?)	(Maria sent **her program** to Greece.)
6c.	
[pço ˈpraɣma tisˈestile i maˈria stin eˈlάða]	[i maˈria ˈestile to ˈ**proɣra**ˈ**ma** tis stin eˈlaða]
(Which thing of hers did Maria send to Greece?)	(Maria sent her **program** to Greece.)
6d.	
[pçaˈnu to ˈproɣrama ˈestile i maˈria stin eˈlaða]	[i maˈria ˈestile to ˈproɣraˈma **tis** stin eˈlaða]
(Whose program did Maria send to Greece?)	(Maria sent **her** program to Greece.)
6e	
[pça ˈestile to ˈproɣraˈma tis stin eˈlaða]	[i **ma**ˈ**ria** ˈestile to ˈproɣraˈma tis stin eˈlaða]
(Who sent her program to Greece?)	(**Maria** sent her program to Greece.)

2.4.2 Results

The set up of the material of the third experiment permits a syntagmatic and a paradigmatic, as well as a contextual comparison of the F0 contour of the paratactic~enclitic structure. Thus in the investigated [to ˈproto ˈɣrama ~ to ˈproɣraˈma tis] opposition, the word stresses of the paratactic structure can be compared with each other (syntagmatic plane), whereas the enclitic's word and enclitic stress has already been covered by experiment II. The paratactic structure's word stresses can be compared with the enclitic's word and enclitic stress (paradigmatic plane), a subject partly covered by experiment I. Finally, the paratactic's word stresses can be compared in different focus productions (contextual plane), a subject partly covered by experiment I, whereas the enclitic's contextual manifestations of the F0 structure is covered by experiment II.

As an example, one speaker's raw data of F0 of the paratactic~enclitic structures [to ˈproto ˈɣrama ~ to ˈproɣraˈma tis] in different contextual frames is shown in Figure 2.13; duration as well as intensity data is shown in Appendix 2.1 and 2.2. Five speaker's mean F0 values and standard deviations are shown in Table 2.7 and Appendix 2.3, their grand mean in Figure 2.14.

In Figure 2.13, a typical F0 division across the utterances is evident, i.e. major F0 variations in the focal part, minor variations in the prefocal part, and (almost) no variations in the postfocal part of the utterances. The paratactic~enclitic noun phrases [to ˈproto ˈɣrama ~ to ˈproɣraˈma tis] appear with the same F0 structure in prefocal (5a~6a) and postfocal (5e~6e) positions, i.e., minor F0 variations for the word and/or enclitic stress prefocally and no variations postfocally (Experiment I & II).

When the whole phrases are in focus (5b~6b), the major F0 variations are on the word stress of [ˈɣrama] and on the enclitic stress of [ˈproɣraˈma] for the paratactic and enclitic structure respectively; the same holds true when the noun [ˈɣrama] and the enclitic [tis] are in focus (5d~6d). However, when the adjective [ˈproto] of the paratactic structure [to ˈproto ˈɣrama] is in focus (5c), the major F0 variation is associated with the word stress of the adjective whereas, in contrast, when the noun [ˈproɣrama] of the enclitic structure [to ˈproɣraˈma tis] is in focus (6c), the major F0 variation is associated with the enclitic stress of the enclitic structure; duration and intensity show larger relative values associated with the major F0 variations within the paratactic ~ enclitic structure.

In the following we shall concentrate on 5a, 5c, 5e ~ 6a, 6c, 6d, which have partly similar and partly different F0-contours, and regard 5b and 5d as the same in prosodic manifestation with 6b, 6d (Figure 2.13).

Prefocal distribution

In prefocal position (Table 2.7, 5a~6a), on the syntagmatic plane, the five speakers' mean F0-contour of the paratactic~enclitic structure [to ˈproto ˈɣrama ~ to ˈproɣraˈma tis] rose from 126 Hz at the beginning of the first word stressed syllable to 156 Hz at the end of it, and from 119 to 147 Hz on the second word stressed syllable of the paratactic structure; the enclitic's word stressed syllable started at 129 Hz at the beginning and rose to 152 Hz at the end of the syllable, and the enclitic stressed syllable started at 122 Hz and rose to 144 Hz. On the paradigmatic plane, the paratactic~enclitic F0 contour was quite the same.

All five speakers produced the majority of the paratactic~enclitic structure with two F0 rises associated with the stressed syllables. However, whereas the first F0 rise was a rather independent tonal gesture, the second F0 rise was tonally assimilated to the upcoming, larger F0 gesture, particularly for speakers 1 and 5. The first and second F0 rises of both members of the pair were not produced in a constant way but some speakers showed a larger first F0 rise and some speakers a larger second one. On the other hand, both F0 rises of the first member were larger than the ones of the second member except for speaker 3 who produced the second stress of both members of the pair with equal F0 rises of 26 Hz (111~137, 114~140 Hz).

Most of the speakers produced the first stressed syllable of each member of the pair with a higher F0 except for speaker 4 who produced the second stressed syllable of the second member with a higher F0 (189~192 Hz) and speaker 5 who produced the second syllable of both members of the pair with a higher F0 (126~132, 125~129 Hz).

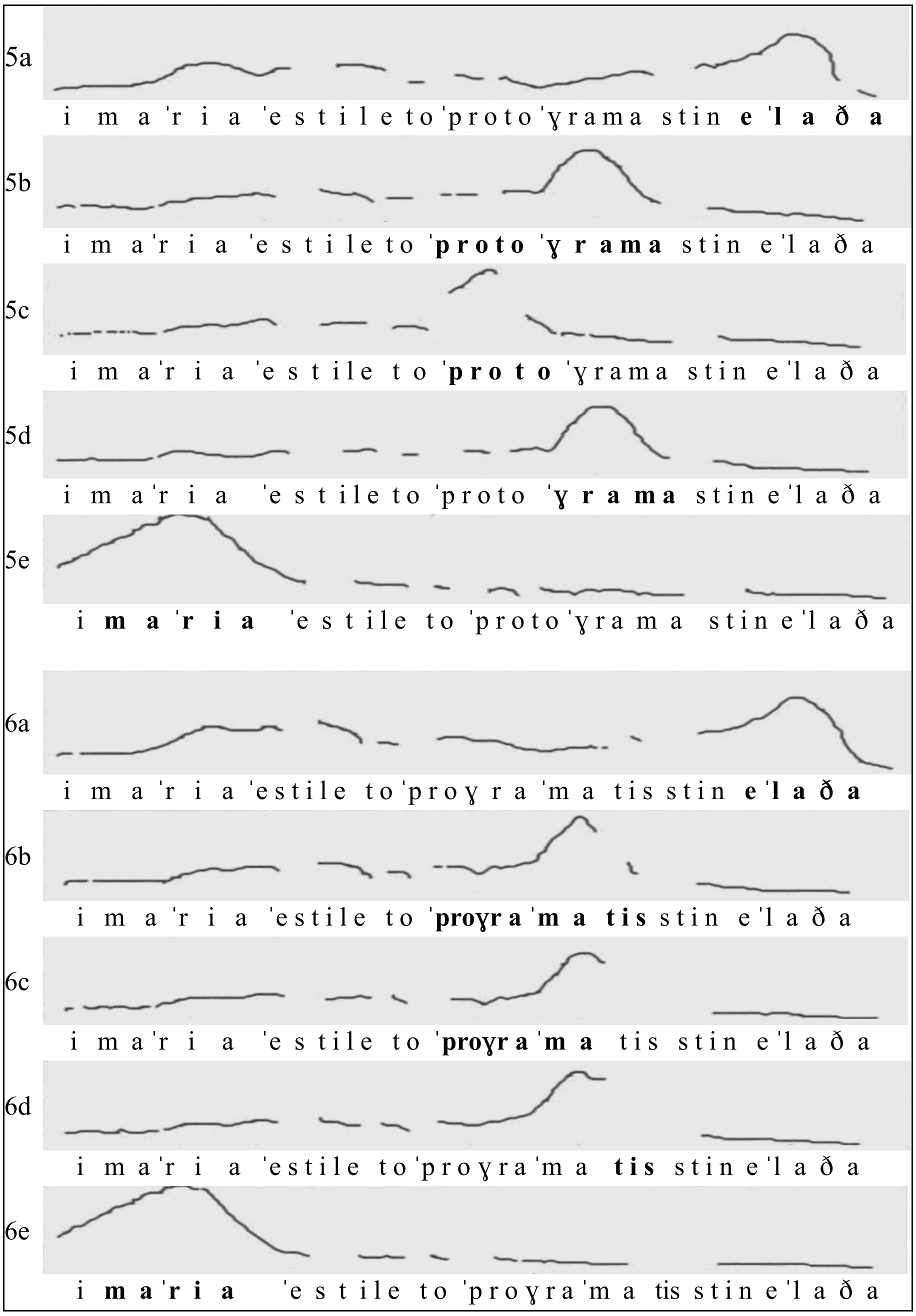


Figure 2.13. One speaker's raw F0 contours of the minimal pair [ˈproto ˈɣrama ~ ˈproɣraˈma tis] (5-6) in the carrier sentence [i maˈria ˈestile to __stin eˈlaða] with different focal organizations (a-e).

Table 2.7. Five speakers' mean voice fundamental frequency (Hz), standard deviation (sd) and grand mean ($\bar{\bar{X}}G$) of the minimal pair [ˈproto ˈɣrama ~ ˈproɣraˈma tis] in the carrier sentence [i maˈria ˈestile to ___ stin eˈlaða] in prefocal (5a~6a), focal (5c~6c) and postfocal position (5e~6e).

(3a~4a)	ˈp r o	t o	ˈɣ r a	m a	ˈp r o	ɣ r a	ˈm a	t i s
1	Hz	102	128	101	114	98	114	91	99
	sd	5.2	4.9	2.6	6.6	1.6	4.4	2.1	1.0
2		148	183	122	156	158	190	133	162
		2.7	2.7	2.7	4.1	5.7	7.0	7.5	5.7
3		124	146	111	137	125	142	114	140
		3.7	3.5	2.1	2.5	2.1	2.5	2.0	3.8
4		148	200	151	197	156	189	161	192
		4.4	3.5	4.1	2.7	4.1	7.4	11.9	9.0
5		111	126	114	132	111	125	114	129
		2.0	3.5	5.6	5.7	2.6	2.1	3.0	1.6
	$\bar{\bar{X}}G$*	126	156	119	147	129	152	122	114
	sd	21.0	33.3	18.9	31.6	26.7	35.6	26.1	34.9
(3b~4b)	**ˈp r o**	**t o**	ˈɣ r a	m a	**ˈp r o**	**ɣ r a**	**ˈm a**	t i s
1	Hz	102	208	95	88	97	103	118	186
	sd	2.7	13.0	5.0	2.7	5.1	10.0	17.5	4.1
2		149	246	107	96	161	198	204	248
		12.4	4.1	7.5	6.5	6.5	9.0	15.5	6.7
3		132	184	81	79	105	142	122	186
		2.5	9.6	1.0	1.0	5.0	2.7	2.5	6.1
4		165	241	132	121	166	200	186	221
		3.5	4.1	2.7	2.2	4.1	6.1	4.1	4.1
5		114	155	79	77	114	131	131	151
		2.5	5.0	1.6	1.6	1.7	1.6	1.6	2.2
	$\bar{\bar{X}}G$	132	206	98	92	128	154	152	198
	sd	25.5	38.4	21.7	17.7	32.4	42.7	39.8	37.1
(3e~4e)	p r o	t o	ˈɣ r a	m a	ˈp r o	ɣ r a	ˈm a	t i s
1	Hz	84	82	81	81	83	82	81	79
	sd	1.3	1.8	1.3	1.0	3.2	2.6	2.6	2.6
2		98	103	94	100	95	94	94	91
		2.7	4.4	4.1	5.0	2.6	3.6	3.7	2.1
3		82	82	81	81	87	88	86	85
		1.7	3.0	2.1	3.8	1.0	1.6	1.6	1.0
4		116	118	115	115	121	122	119	119
		2.2	2.7	1.0	2.8	4.7	3.3	3.4	3.1
5		87	85	84	82	85	85	84	82
		1.6	1.6	1.4	1.6	3.3	2.6	1.6	1.6
	$\bar{\bar{X}}G$	93	94	91	91	94	94	92	91
	sd	14.0	16.0	14.4	15.2	15.6	16.1	15.4	16.1

Focal distribution

In focal position (Table 2.7, 5c~6c), on the syntagmatic plane, when the paratactic's adjective [ˈproto] is in focus, the first word stressed syllable rose from 132 to 206 Hz, a variation of 74 Hz, and the second word stressed syllable fell from 98 to 92 Hz. The enclitic's word stressed syllable had an F0 rise of 26 Hz, from 128 to 154 and the enclitic stressed syllable an F0 rise of 46 Hz, from 152 to 198 Hz. On the paradigmatic plane, the paratactic's first word stressed syllable had an extended F0 rise of 74 Hz in comparison with the enclitic's 26 Hz F0-rise whereas the paratactic's second word stressed syllable had a fall of 6 Hz (98~92) as a part of the postfocal F0 fall in comparison with the enclitic's 46 Hz F0 rise.

All five speakers produced the paratactic structure with one F0 rise associated with the first word stressed syllable in focus whereas the postfocal, second word stressed syllable showed a falling and flattening F0 contour at a low level.

On the other hand, the enclitic structure appeared with two F0 rises, a word stress rise, except for speaker 1 who assimilated the word stress rise to the upcoming focal rise with a slight rising and flattening F0 contour (97~103 Hz), and a focal, enclitic stress rise. Moreover, the speakers showed a tendency to produce the focal rise in an anathetic structure (upstepping), particularly speakers 2 and 5 whose focal rise started at about the same level the word stress rise reached (Sp. 2 198~204, Sp. 5 131~131 Hz).

Both the F0 maximum and the focal rise of the enclitic structure were larger than the ones of the word stress whereas the corresponding values of the paratactic~enclitic structure were not constant among the speakers.

Postfocal distribution

In postfocal position (Table 2.7, 5e~6e), the F0 contour of the paratactic~enclitic structure was quite the same, declining slightly with an intra-variation of 3 Hz and an inter-variation of 1 Hz. Speakers 1 and 5 produced both structures with a slight declining F0 contour at a low level. Speaker 2 produced both word stressed syllables of the paratactic structure with small F0 rises of 5 Hz for the first word stressed syllables (98~103 Hz) and 6 Hz for the second word stressed syllables (94~100 Hz) and produced hardly any F0 rises for the enclitic structure. Speakers 3 and 4 produced both structures with negligible F0 variations, speaker 3 the word stressed syllable of the enclitic structure (116~118 Hz), speaker 4 the word stressed syllable of the paratactic structure 121~122 Hz).

Contextual distribution

On the contextual plane (Figure 2.14), the paratactic's first word stressed syllable in focus had an extended F0 rise of 74 Hz (132~206) in comparison with the prefocal one which had a rise of 30 Hz (126~156) whereas the post focal one had a rather level F0 contour (93~94 Hz); the second word stressed syllable had a declining fall of 6 Hz (98~92) in comparison with the prefocal one which had an F0 rise of 27 Hz (119~147) and the postfocal one which had a rather level F0 contour as well (91~91 Hz). The enclitic's word stressed syllable in focus had an F0 rise of 26 Hz (128~154), the prefocal one 23 Hz (129~152) whereas the postfocal one was rather level (94~94 Hz); the enclitic stressed syllable had an extended F0 rise of 46 Hz (152~198) in comparison with 22 Hz rise of the prefocal (122~144 Hz) and a declining 1 Hz (92~91) of the postfocal one.

All five speakers produced the focal word stressed syllables of the paratactic structure as well as the enclitic stressed syllables of the enclitic structure with the largest F0 rise, the highest F0-top and the largest F0-fall within the utterances in a constant way. Apart from the larger F0-rises, the focal stress syllables showed a rather anathetic F0-contour in comparison to the prefocal stress syllables which appeared in an alternating F0-contour.

The F0-rises of the focal stress syllables of the paratactic structure started at a higher F0-level than the prefocal ones for speakers 3 and 4 (Sp. 3 124~132, Sp. 4 148-165 Hz) whereas speakers 1, 2 and 5 started their F0-rises at about the same level for the two positions. On the other hand, when the enclitic structure was in focus, the enclitic stress syllables started their F0-rises higher than the prefocal ones whereas the word stress syllables had considerable interspeaker variation, especially speaker 3 who produced the prefocal rises at a higher level than the focal ones (125~105 Hz).

The postfocal F0 variations were limited to the minimum at the lowest level within the utterances in a constant way for all speakers.

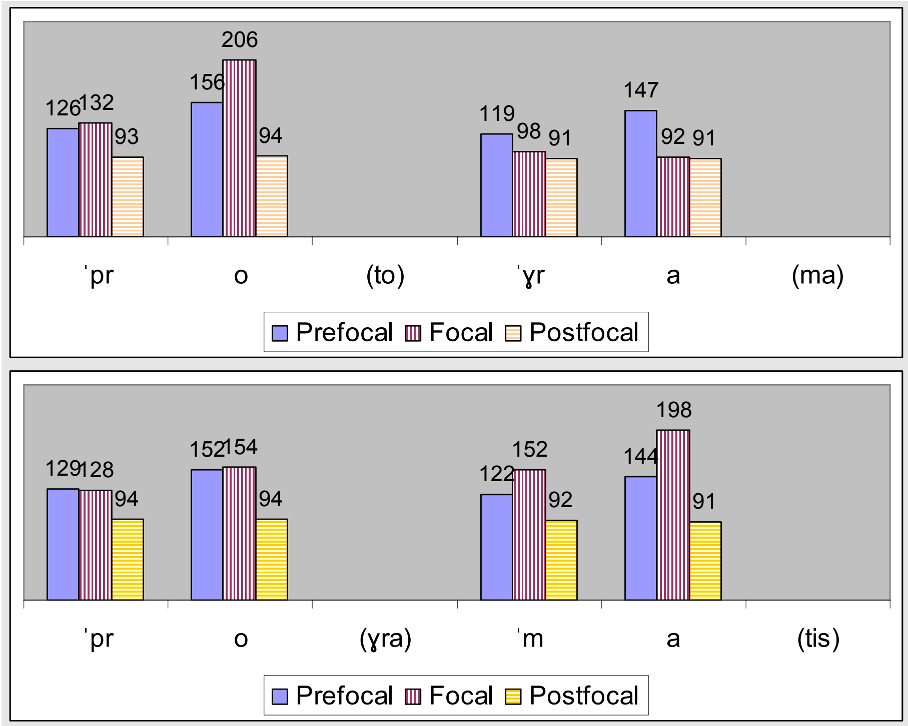


Figure 2.14. Five speakers' average F0 (grand mean) in Hz of the minimal pair [ˈproto ˈɣrama ~ ˈproɣraˈma tis] in the carrier sentence [i maˈria ˈestile to ___ stin eˈlaða] in prefocal, focal and postfocal position.

2.4.3 Discussion

In this experiment the prosodic manifestation of a paratactic structure was compared to the prosodic manifestation of an enclitic structure. The paratactic structure is composed of two word stresses in a noun phrase whereas the enclitic structure is composed of one word stress and one enclitic stress also in a noun phrase.

In prefocal and postfocal positions the F0 contour of the paratactic~ enclitic structure appears quite the same, i.e. with no apparent indication that in the paratactic structure we have two word stresses whereas in the enclitic structure we have one word stress and one enclitic stress; in prefocal position, the F0 contour has a more or less compressed rise associated with the word and/or the enclitic stress whereas the postfocal F0 contour is low and flattened down with no indication for the word stress ~ enclitic stress distinction either, as already shown in the previous acoustic experiments.

In focal position, the paratactic structure appears with the same F0 contour whether the whole noun phrase or only the noun is in focus, i.e. we have F0 neutralization between narrow ~ broad focus whereas when the adjective is in focus, sentence stress with its acoustic manifestation is applied to the adjective leaving outside its scope the rest of the noun phrase. Thus, the noun phrase is not focused as a whole but rather in accordance with the lexical elements composing it. On the other hand, no matter which element of the enclitic structure is in focus, the focal F0 gesture coincides with the enclitic's enclitic stress, i.e. in the enclitic structure with enclitic stress, the noun phrase is focused as a whole with no possibility of applying sentence stress on the enclitic's lexical element carrying word stress.

Thus, the apparently prosodically similar paratactic ~ enclitic structures may take completely different prosodic manifestations when focused, the enclitic structure imposing a major syntactic constraint to the application of sentence stress whereas the paratactic structure exempts it.

Current theories on the relation between prosody and syntax/semantics have been outlined in Rossi et al. (1981), Rossi (1985). To summarize the three main concepts on the subject: First, there is a direct relation between prosody and syntax, the prosodic structure corresponding to the syntactic one (Lea 1977, Di Cristo 1981a, 1981b, Paccia-Cooper and Cooper 1981). Second, prosody is quite independent with its own structure which, as a rule, does not reflect the syntactic structure (Thorsen 1980, 1983). Third, prosody and syntax do not run parallel to each other but are interrelated in a rather complicated way (Bannert and Thorsen 1986), along with the semantic representation of the utterance (Rossi 1981, 1985).

At a closer inspection of the paratactic ~ enclitic contours, Figure 2.13, we see different F0 realizations for each structure, as a result of the different contextual organizations; on the other hand, different contextual structures are neutralized and appear with the same F0 contour, an interaction of the syntactic structure and prosody. Thus far we are in agreement with Rossi (1985) that both the syntactic and contextual planes contribute to prosody.

However, sentence stress seems to be the milestone of the organization of prosody in Greek for whose application, apart from the focus ~ presupposition division, both the syntactic and morpholexical structures of the sentence are involved since sentence stress is neutralized in lexical elements within an enclitic structure with enclitic stress as described in the last two experiments and, in turn, enclitic stress is the product of the morpholexical and syntactic structures, as outlined in Part 1 of this study.

Thus, although the contextual structure of the sentences may be enough for the right application of sentence stress in simple sentences as the ones used in the first acoustic experiment, in more complicated sentences the lower levels of representation must also be taken into account.

Once sentence stress has been applied, there is hardly any evidence for a further contribution of the contextual and/or syntactic structures of the sentence on prosody, at least for one clause sentences, since narrow ~ broad focus have been found to be neutralized and simple proclitic ~ enclitic structures may not be perceptually distinct as mentioned in note 1, experiment II.

To summarize our point of view, prosody seems to be interrelated with the different levels of the language but once the constraints of these levels have been taken into account prosody may appear quite independent, without any traces between words and minor or major syntactic boundaries.

3 Physiological Study

3.0 Introduction

This study deals with variations in subglottal pressure associated with word and sentence stress in Greek. We have shown previously (Botinis 1982, part 2), that the acoustic parameters of duration, voice fundamental frequency (F0) and intensity contribute in different ways to the manifestation of word and sentence stress in Greek. All three of these acoustic parameters may contribute to word stress, but only duration and intensity may contribute after focus. On the other hand, relative F0 and intensity as well as, to some extent, duration contribute to sentence stress.

The study raises the following questions: (1) Do we have variations of subglottal pressure (Ps) associated with stressed ~ unstressed syllables? (2) If so, which of the acoustic parameters co-vary with subglottal pressure? (3) Do variations in Ps affect one or both of the acoustic parameters of intensity and F0?

Stress has traditionally been described either on an articulatory basis with greater articulatory effort on the part of the speaker (Jones 1950) and/or on an auditory basis with greater loudness on the part of the listener (Bloomfield 1933)

Articulatory effort (i.e. speech effort) may be reflected by the activity of both the laryngeal and respiratory muscles. This has a direct effect on the vibration of the vocal folds for the former and on the subglottal pressure for the latter. An increased tension of the vocal folds will produce a higher fundamental frequency (Ohala 1978) with a perceptual effect of higher pitch as well as greater loudness. Increased subglottal pressure will produce a higher intensity (Ishikki 1964), which is most directly related to loudness, and will also result in a higher fundamental frequency (Ladefoged and Mckinney 1963).

3.1 Experimental Design

3.1.1 Method

An indirect method for estimating Ps from records of oral pressure (Po) has been applied (Holmberg 1980, Smitheran and Hixon 1981), and its validity has been empirically tested (Lofqvist, Carlborg and Kitzing 1982). This method exploits the fact that during the production of voiceless stops Po and Ps are identical. By constructing a suitable linguistic material where voiceless stop consonants and vowels alternate, it is possible to obtain indirect measurements of Ps. This is done by linearly interpolating the Po records of the stop consonants and thus estimating the Ps associated with the intervening vowels.

3.1.2 Subjects

The main subjects of this experiment are two male students, in their mid twenties, brought up and educated in Athens; in addition, the present investigator, a male in his early thirties at the time of the experiment was recorded. None of the speakers has had any known history of speech, hearing, neurological, or respiratory disorders. All three speakers are monolinguals and speak Standard Athenian.

3.1.3 Physiological Analysis

The recordings took place in a sound-treated room at the Phonetics Laboratory, Lund University. Oral pressure was sampled through a plastic tube, 10 cm long and with an inner diameter of 2 mm. The tube was held between the lips, and the open end of the tube was positioned just behind the upper teeth. The tube was coupled to a differential pressure transducer. After suitable amplification, the pressure signal was recorded on one channel of a 3968:A multichannel instrumentation recorder (Hewlett Packard). During the experiment, the pressure signal was also monitored on an oscilloscope in order to detect clogging of the tube. Static calibration of the recording system was performed before and after the recording session using a water manometer. Oscillograms of the test sentences from the tape recordings were made. For the segmentation of the utterances, a duplex oscillogram and the Po contour were used. Po was measured in cm H_2O on a linear scale.

3.1.4 Procedure

Proper meaningful sentences were set up for the purpose of this physiological experiment. The subjects were provided with a set of paper cards with the test material of the experiment, each card containing only one sentence. The test sentences were written with a typewriter, in standard Greek orthography. The subjects were required to read the cards five times each, each time in a different random order, as in everyday speech, with normal tempo and loudness. The experiment took place on one occasion, lasted about ten minutes, and the subjects performed their task relatively easily, the tube in the mouth not causing any discomfort.

3.1.5 Speech Material

Two meaningful Greek sentences were set up containing the prosodic minimal pair under investigation [ˈpapa~paˈpa] (pope~priest) in the carrier sentence [ˈɣrapse ___ paˈdu] (write ___ everywhere). For the purpose of this experiment, the segmental structure of the minimal pair was composed of an alternation of stops and vowels (see method). Both sentences were imperative with the same syntactic structure, i.e. VP-NP-AP.

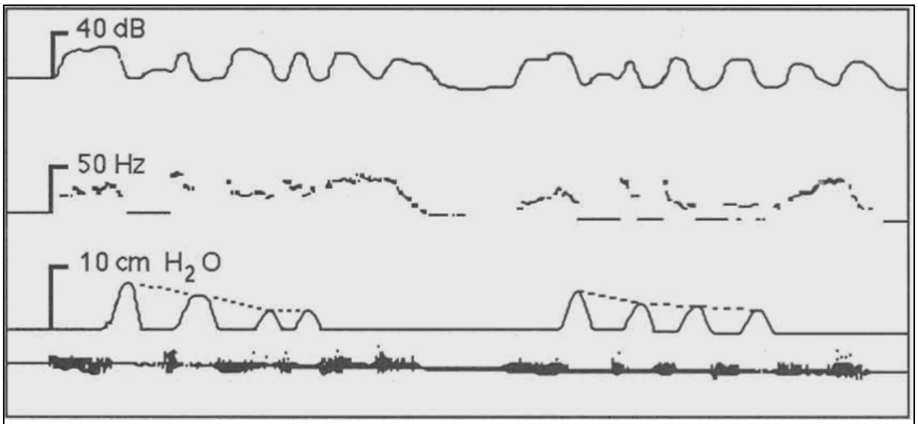
Each of the two test sentences was the carrier of four different contextual frames listed in Table ΠΙ.Ι (bold letters indicate focus). First, they were pronounced neutrally, i.e. with no contextual information, productions la~2a. By having the minimal pair in the middle position we could easily and accurately examine its manifestation in postfocal, focal and prefocal position by shifting sentence stress to initial, medial and final positions respectively. This was done by formulating appropriate questions to make the speaker choose one of the elements of the sentence as the focus and the carrier of the information required by the question. Thus, in productions lb~2b the verb [ˈɣrapse] is in focus having [ˈpapa~paˈpa] in postfocal position, in lc~2c [ˈpapa~paˈpa] is in focus and in ld~2d the adverb [paˈdu] is in focus, having [ˈpapa~paˈpa] in prefocal position.

Table III.I. The minimal pair [ˈpapa~paˈpa] in the carrier sentence [ˈɣrapse ___ paˈdu] with different contextual organizations (see text).

Contextual frame	Test sentences
1a.	
None	[ˈɣrapse ˈpapa paˈdu]
	(Write pope everywhere.)
1b.	
[ti na ˈkano ton ˈpapa paˈdu]	[ˈ**ɣrapse** ˈpapa paˈdu]
(What shall I do with pope everywhere?)	(**Write** pope everywhere.)
1c.	
[ti na ˈɣrapso paˈdu]	[ˈɣrapse ˈ**papa** paˈdu]
(What shall I write everywhere?)	(Write **pope** everywhere.)
1d.	
[pu na ˈɣrapso ˈpapa]	[ˈɣrapse ˈpapa **pa**ˈ**du**]
(Where shall I write pope?)	(Write pope **everywhere**.)
1a.	
None	[ˈɣrapse paˈpa paˈdu]
	(Write priest everywhere.)
1b.	
[ti na ˈkano ton paˈpa paˈdu]	[ˈ**ɣrapse** paˈpa paˈdu]
(What shall I do with priest everywhere?)	(**Write** priest everywhere.)
1c.	
[ti na ˈɣrapso paˈdu]	[ˈɣrapse **pa**ˈ**pa** paˈdu]
(What shall I write everywhere?)	(Write **priest** everywhere.)
1d.	
[pu na ˈɣrapso paˈpa]	[ˈɣrapse paˈpa **pa**ˈ**du**]
(Where shall I write priest?)	(Write priest **everywhere**.)

3.2 Results

One speaker's pressure variations of the minimal pair [ˈpapa~paˈpa] in their carrier sentences, in neutral production, as well as the acoustic parameters of duration, F0, and intensity are shown in Figure 3.1. In the first utterance (left), with stress on the first syllable of the first member of the pair, the Ps is at a high level at the very beginning of the syllable, i.e. during [p] and then falls gradually to the beginning of the next syllable where it remains level. In the second utterance (right), with stress on the second syllable of the second



Time (ms) |-------| 200

Figure 3.1. One speaker's intensity, F0, oral pressure, and oscillogram (from top to bottom) of the minimal pair [ˈpapa~paˈpa] (left~right), in the carrier sentence [ˈɣrapse ___ paˈdu]. The interpolated broken line connecting the oral pressure of voiceless stops is an estimation of subglottal pressure during vowels.

member of the pair, the Ps does not change but remains level during both syllables. In both utterances, the stressed syllables are longer than the unstressed ones and the Ps declines from the beginning to the end.

In Figure 3.2, the oral pressure variations of the three productions of the minimal pair [ˈpapa~paˈpa] in postfocal, focal, and prefocal position for all three speakers is shown. The first speaker produces the stressed syllable of the first member of the pair with about a 2.5 cm H_2O higher Ps than the unstressed syllable and with equal Ps for both unstressed and stressed syllables of the second member of the pair in postfocal position. In focal position, the stressed~unstressed difference for the first member of the pair is enlarged by about 1.5 cm H2O whereas, for the second member of the pair, the stressed syllable appears with about a 2.5 cm H2O higher Ps than the unstressed one. In prefocal position, the minimal pair shows the same degree of Ps variation between the stressed~unstressed syllables as in postfocal position.

The second speaker produces the stressed syllable of the first member of the pair with higher Ps in postfocal and focal positions (up to 5 cm H_2O), but with lower Ps than the unstressed syllable in prefocal position (≈2 cm H_2O); for the second member of the pair, the Ps appears lower for the stressed syllable in postfocal position (≈2.5 cm H2O) but higher in focal (≈5 cm H_2O) and prefocal (≈2 cm H_2O) positions.

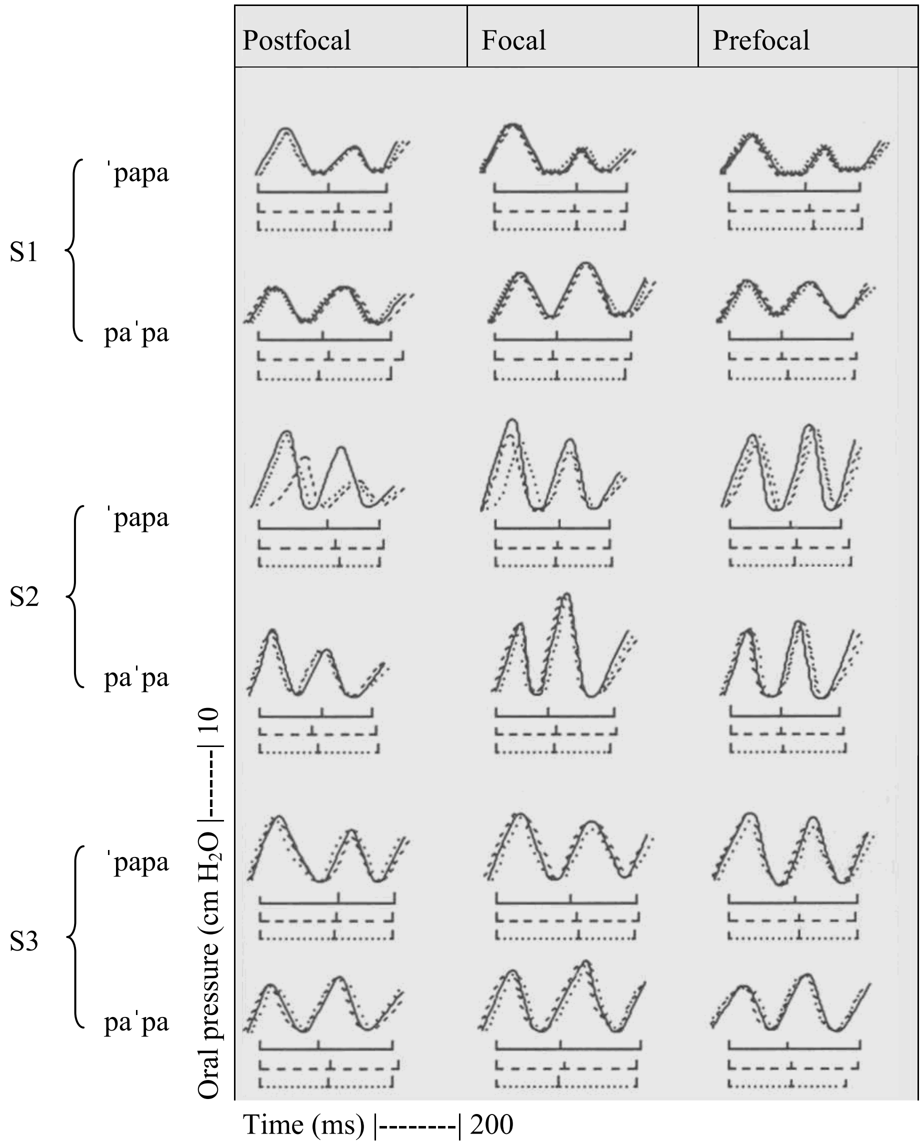


Figure 3.2. Oral pressure and syllabic durations of three speakers' (S1, S2, S3) three productions (solid, broken line, dots) of the minimal pair [ˈpapa~paˈpa] in the carrier sentence [ˈɣrapse ___ paˈdu] in postfocal, focal, and prefocal position.

The third speaker produced the stressed syllables of both members of the pair in all positions with higher Ps (up to 3 cm H_2O) than the unstressed ones and with higher Ps for the second member of the pair in focal position.

To summarize, for stressed~unstressed syllables in focus, the stressed syllables have a greater Ps than the unstressed ones. The same holds for the stressed~unstressed syllables out of focus (with some reservations for the second speaker) on the syntagmatic and/or the paradigmatic plane.

3.3 Discussion

Recent studies on the aerodynamics of stress have pointed out the direct relation between Ps, Po and oral air flow (Vo) which corresponds to Po variations (Stathopoulos and Weisner 1985a). A direct relation between Ps and intensity is, by and large, well accepted (Ishikki 1964, Stathopoulos and Weisner 1985b, Tanaka and Gould 1983), although the aerodynamic power (mean flow rate times intrapulmonic pressure) may appear to be an even better correlate of intensity (Tanaka and Gould 1983).

The analysed Greek data show that the stressed syllables either in focus or out of focus are associated with increased Ps as a rule and that there is a declination of Ps from the beginning to the end of the utterances (Lieberman 1967). Moreover, the stressed syllables may appear with the same Ps as the unstressed ones, especially when the former are preceded by the latter, an indication that the perceptual system may compensate for the declination effect (Pierrehumbert 1979).

Our results agree with Ladefoged's (1967) reports of an increased activity of the respiratory muscles on stressed syllables, whether emphatically stressed or not, as well as with Stathopoulos and Weismer (1985b) who report Vo variations associated with the stressed~unstressed syllables, to the degree Vo corresponds to Po and Ps under different experimental conditions. However, van Katwijk (1974) and Ohala (1977) report increased expiratory activity only on emphatically, and not on normally stressed syllables.

If we combine the results reported in Botinis (1982, part two) with the results obtained in this study, we see that a stressed syllable has a higher relative intensity associated with a higher Ps (Ladefoged and Mckinney 1963, Isshiki 1964). The prefocal and focal stressed syllables may have higher F0, longer duration and higher intensity; in these positions, it is difficult to examine the influence that Ps may have on the acoustic parameters of F0 and intensity. If we consider the postfocal stressed syllables, we see that the higher Ps has an effect only on intensity, whereas

F0 remains low and flat (part two, part four). This has interesting theoretical implications. First, it implies that the human speech mechanism does not produce a quasi-constant Ps, but a varied one that can produce linguistic effects. Second, variations in Ps do not necessarily have an effect on F0.

Studies reviewed in Ohala (1978) argue that F0 is mainly regulated by the laryngeal muscles. Moreover, both F0 and intensity (the latter being correlated to Ps) are supposed to be influenced by the larynx. Thus, according to Ohala, not only are observed F0 variations independent of Ps variations by and large but, to some extent, Ps variations may be caused by laryngeal activity.

The Greek data agree with Ohala's first point; although there is no EMG data in this experiment, there is no evidence that the subglottal system affects the F0 contour of the postfocal word stress, i.e. the laryngeal muscles must be responsible for the F0 structure. In this position, the Ps for a stressed syllable is usually higher; assuming that the vocal cords maintained a constant level of tension, the F0 contour should automatically rise higher because of the higher Ps. But this is not the case. The structure of the language requires that the F0 contour be flat in postfocal position. Accordingly, the laryngeal muscles will have to keep F0 flat, although the Ps variations should have an effect on F0.

> Unless an adjustment is provided in the tension of the vocal folds, increased subglottal pressure results automatically in an increased rate of vocal fold vibration. (Lehiste 1970, p. 125).

We disagree with Ohala's second point concerning the same postfocal position. Here there are no F0 variations caused by increased laryngeal activity that could raise the Ps. Moreover, it is clear at this point that the Greek speakers do not produce a constant Ps but a varied one. Since the focal organization of a sentence is so variable and common in speech with different prosodic manifestations across the utterance, one can hardly expect good correlations between Ps and F0 (Collier 1975).

Lieberman (1967) raised the question whether the observed F0 variations were accomplished by the activity of the laryngeal muscles or the respiratory system. Lieberman had records only on Ps and assumed that the laryngeal muscles maintained a constant level of tension. This was thought to be the case for declarative sentences in English. In yes~no questions, where there is a final F0 rise without any variation in Ps, Lieberman assumed the laryngeal muscles to be active. Liberman's hypothesis that Ps has a causal effect on F0 does not agree with the Greek data reported here. In postfocal position, where stress is associated with Ps variations, flat F0 is the rule, although minor F0 variations, up to 10 Hz, may occasionally be found.

Atkinson (1978) argues that different physiological mechanisms may be involved at different F0 levels. In particular, in declarative sentences and for low F0 values (80-100 Hz), F0 is mostly governed by Ps. However, in the analysed Greek utterances where F0 is within or very close to the 80-100 Hz low F0-level, i.e. in postfocal position, Ps still does not have the expected effect on F0, i.e. 3-7 Hz/cm H_2O (Baer 1979). My interpretation of the Ps data as to whether it is the larynx or lungs that controls F0 (pitch), the answer is the larynx (Vanderslice 1967, Ohala 1977).

Of course variations in Ps can influence F0, other things being equal, but F0 must take the configuration required by the structure of a particular language. In the case of the investigated Greek data, F0 serves one purpose (semantic weighting) and Ps another (rhythmic structuring). It is possible that a stressed syllable may gather all acoustic and physiological parameters, but it is far from reality that every stressed syllable is associated with an F0 (pitch) change (Hyman 1977), (see part 2).

In Greek, there are two facts that suggest that the acoustic parameters of F0 and intensity are independent of each other and not produced by the same mechanism. First, heightened Ps in postfocal position has not the expected effect on the F0-contour. Second, in postfocal position, intensity is combined with duration to give the impression of word stress, whereas in focal position intensity is combined with F0, and to a variable degree with duration, for the impression of sentence stress.

In Swedish, F0 variations are distinctive both at word and sentence level, though F0 spans a smaller range after focus (Bruce 1977). In a general study of the Scandinavian languages (Gårding 1977a), the F0 manifestations were reported to be different for the "same" tonal phenomena. Even for the Swedish dialects the F0 contour may exhibit a great variability depending on the dialect. Another comparison of the intonational patterns of Swedish, Greek, and French (Gårding et al. 1982) demonstrated how the three languages utilize F0 in different ways for the "same" prosodic categories, i.e. word stress, sentence stress, etc. These observations suggest the large interdialectal variability speakers may have when producing the same linguistic category either in one or across languages.

4 Perceptual Study

4.0 Introduction

In the perceptual part of this study six experiments were carried out, the first four experiments on word stress perception and the last two ones on enclitic stress perception. The purpose of the first experiment was to investigate which of the acoustic parameters contributes most to the perception of word stress after focus. The second experiment was to investigate which of the acoustic parameters contributes most to word stress perception before focus. The third experiment was to investigate the perceptual relevance of F0 in the frequency dimension and the fourth experiment in the time dimension for the word stress distinction. The fifth experiment was to investigate which of the acoustic parameters contributes most to the perception of enclitic stress and the sixth experiment was to investigate the perceptual relevance of F0 for the enclitic stress distinction.

4.1 Experimental Design

4.1.1 Subjects

The speaker of the speech material is a male, 35 years old at the time of the experiments, brought up and educated in Athens; he speaks what is considered to be standard Athenian. For each experiment, ten different volunteer high school students between 17 and 18 years old, both male and female, having no hearing, neurological or respiratory disorders took part. All listeners have grown up and have been educated in the same educational section in Athens where the speaker received his education. All listeners were students in the same school and speak about the same sociolect.

4.1.2 Procedure

The experiments were conducted at the listeners' school, the 2nd high school, Athens. Listeners were told that the experiments were dealing with speech perception and the sentences they were going to hear were produced by a computer. Their task was to mark what they perceived as a stressed syllable with an acute accent according to the Greek orthography; in cases they were uncertain about they would mark the syllable closest to their preference. The answer sheet was a piece of paper with ten rows, each row showing the synthetic stimuli in a random order. The time between the stimuli was two seconds and between the rows five seconds. Each experiment lasted about a quarter of an hour and the listeners said their task was an easy one, although sometimes they were uncertain; they said would like to participate in other experiments as well.

4.1.3 Perceptual Analysis

Synthesis guided by analysis is a powerful tool by which the relative contribution of the acoustic features to the perception of phonetic concepts can be investigated. For this purpose, meaningful Greek sentences were set up and recorded on a Studer A:62 tape recorder (ips 7.5) in a sound-treated room at the phonetics laboratory, Lund University. The frequency response of the tape recorder was flat within ±2 dB from 30 to 14 000 Hz and the signal to noise ratio was 63 dB. The microphone was flat within the frequencies 35 to 17 000 Hz. With the ILS (API program) and some programs developed at the department the test sentences were digitized, analyzed, manipulated and synthesized in a VAX 11/730 computer system.

The synthetic stimuli were of reference ~ target type, i.e. the acoustic manipulations were restricted to the acoustic manifestation of the speech material under investigation with reference to the original sentences, either in one or in several successive steps. The produced synthetic stimuli were recorded and the experimental tapes were presented to the listeners binaurally through headphones on a Revox A:77 tape recorder.

4.2 Perception of Word Stress

4.2.0 Introduction

In acoustic experiment I, intensity and duration were found to correlate with word stress after focus, intensity, duration and F0 before focus. In the four perceptual experiments on word stress to follow, the perceptual relevance of the above acoustic correlates was investigated.

The main purpose of perceptual experiment I was to find out which of the two acoustic correlates, intensity or duration, is the stronger perceptual cue and to what degree to the perception of word stress after focus; is only one of the above correlates able to convey the concept of postfocal word stress or is the combination of both an absolute necessity? Moreover, the influence of the context on postfocal word stress perception was investigated. Last, we could see whether listeners were able to perceive stress distinctions after focus where the acoustic parameters are weakly manifested and especially with synthetic speech which is degraded in quality. In this experiment, to exclude other acoustic parameterσ but intensity and duration as possible perceptual cueσ, the F0 and the formant structure of the vowels composing the prosodic minimal pair under investigation were examined as well.

In experiment Π, we investigated the relative contribution of intensity, duration and F0 to the perception of word stress before focus. In this experiment we tried to answer three basic questions. First, which is the strongest perceptual cue for word stress in prefocal position? Second, is only one of the acoustic parameters able to convey word stress or is a combination of more acoustic parameters an absolute necessity? Third, can we establish a hierarchy for the acoustic parameters' degree of contribution to the prefocal word stress distinction?

In experiment ΠI, led by the results of perceptual experiment II, we investigated the perceptually relevant point of the F0 contour for the identification of prefocal word stress, testing synthetic stimuli in successive steps in the frequency dimension whereas, in experiment IV, we investigated the perceptual relevance of the timing of the F0 contour, testing synthetic stimuli in successive steps in the time dimension in prefocal position.

Recent studies on the acoustics of stress have emphasized the close relation between stress and F0 change; Hyman (1977) and Thorsen (1982) report a number of languages where F0 change is used as a cue to linguistic stress. In a series of experiments, Fry has shown that duration prevails over intensity for stress perception in English (1955) and F0 may outweigh the duration cue producing an all-or-none effect (1958), the formant structure (phonetic quality) of the vowel may be even less effective than the intensity cue (1965). Fry's results were corroborated by Bolinger (1958), Morton and Jassem (1965), although Bolinger questioned the importance of intensity as a perceptual cue.

The relative contribution of the acoustic parameters to stress perception has urged some researches to establish a hierarchy with F0 as the most important cue followed by duration and intensity; Hyman (1977) even argues for the universality of this ranking order. Following the tradition on the hierarchical nature of the acoustic cues on stress Berinstein (1979, p. 2) puts forward the following hypothesis:

> Change in F0, increased duration, and increased intensity, in that order, constitute the unmarked universal hierarchy for perception of stress in languages with no phonetic contrasts in tone or vowel length; in languages with such contrasts the perceptual cue correlated with that contrast i.e. F0 with tone and duration with length will be superseded by the other cues in the hierarchy.

Thus, according to Berinstein, in Greek which lacks phonemic contrasts in both tone and vowel length, stress perception would be subjected to the unmarked universal hierarchy principle with F0 prevailing over duration and duration over intensity. Although Greek has not been subjected to perceptual experimentation up to now, it has been shown (Botinis 1982) that it is the prosodic and semanticontextual structure of the utterance which determines the relative contribution of the acoustic parameters to stress distinction.

In a comparative study of English and Japanese, Beckman (1986) has shown that Japanese, an accent (non-stress accent) language, uses F0 to a higher degree in relation to the other prosodic correlates than does English, a stress (stress accent) language, for the accent and stress distinction respectively (see part 1). Moreover, in English, the combination of both intensity and duration into a single prosodic feature, the energy integral (total amplitude), is more closely associated with stress than is F0, in both production and perception.

4.2.1 Perceptual Cues of Postfocal Word Stress

4.2.1.1 Description of Stimuli

Ten synthetic stimuli were set up for this experiment (Table 4.1). The listeners were asked to identify the investigated words as [ˈnomo] (law) or [noˈmo] (county) and mark with an acute accent the grapheme {nomo} presented to them on an answer sheet of paper with ten rows, each row containing the ten stimuli in random order.

The first two stimuli were the two human produced sentences [i **maˈria** ˈiksere to ˈnomo kaˈla] (**Maria** knew the law well) (Figure 4.1), and [i **maˈria** ˈiksere to noˈmo kaˈla (**Maria** knew the county well) (Figure 4.2). The two sentences were answers to the question [pça ˈiksere to ˈnomo~noˈmo kaˈla] (who knew the law ~ county well?), i.e. the word [maˈria] was in focus. The two test sentences were digitized, analyzed, and then synthesized. Both sentences appeared with the regular acoustic manifestation (part two), the minimal pair [ˈnomo~noˈmo] showing both intensity and duration contributions on the syntagmatic and/or the paradigmatic planes for the word stress distinction.

In stimulus 3, the word [ˈnomo] was transferred from its natural context to [noˈmo]'s context and in stimulus 4 we the words [noˈmo] and [ˈnomo] were transferred to each other's context. In stimulus 5, using the facilities of the computer system, we expanded the acoustical signal of the minimal pair [ˈnomo~noˈmo] and changed the intensity envelope of [ˈnomo] (no=-37, mo=-42 dB) to that of [noˈmo] (no=-38, mo=-39 dB), (Figure 4.3). In stimulus 6 we changed the four segments' duration of [ˈnomo] (n=57, 6=70, m=70, o=64 ms) to the four segments' duration of [noˈmo] (n=51, o=64, m=70, o=96 ms) (Figure 4.3). Stimulus 7 was a combination of stimuli 5 and 6, i.e. we changed the intensity envelope and segmental duration of [ˈnomo] to that of [noˈmo]. Stimuli 8, 9, and 10 were mirror images of stimuli 5, 6, and 7, i.e., this time we started from [noˈmo] and went to [ˈnomo] changing first only intensity (St.8), then only duration (St.9), and finally both intensity and duration (St. 10). The stimuli manipulations were restricted to the minimal pair [ˈnomo~noˈmo] leaving the context unaffected.

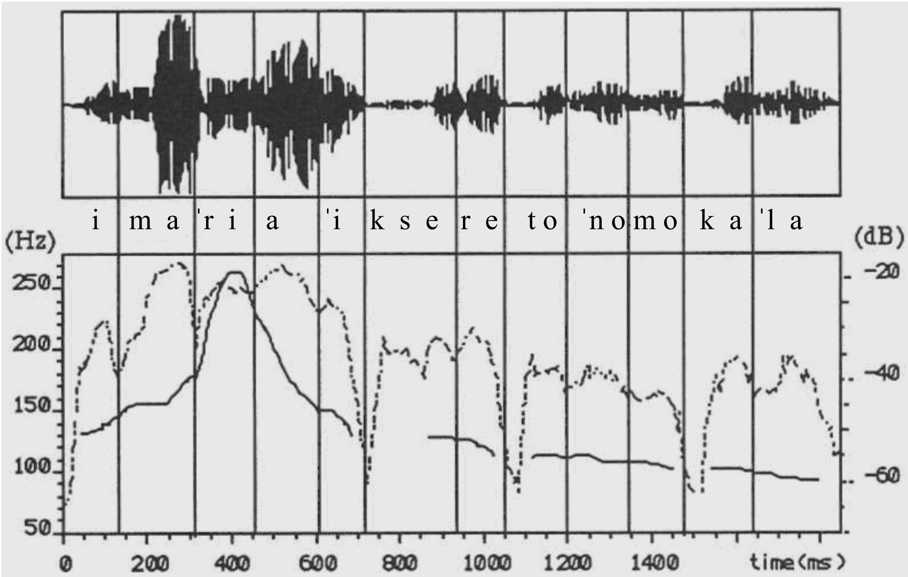


Figure 4.1. Waveform (over), F0 on the left axis (solid line) and intensity on the right axis (broken line) of the first member of the minimal pair [ˈnomo~noˈmo], in the carrier sentence [i maˈria ˈiksere to ___ kaˈla] in postfocal position.

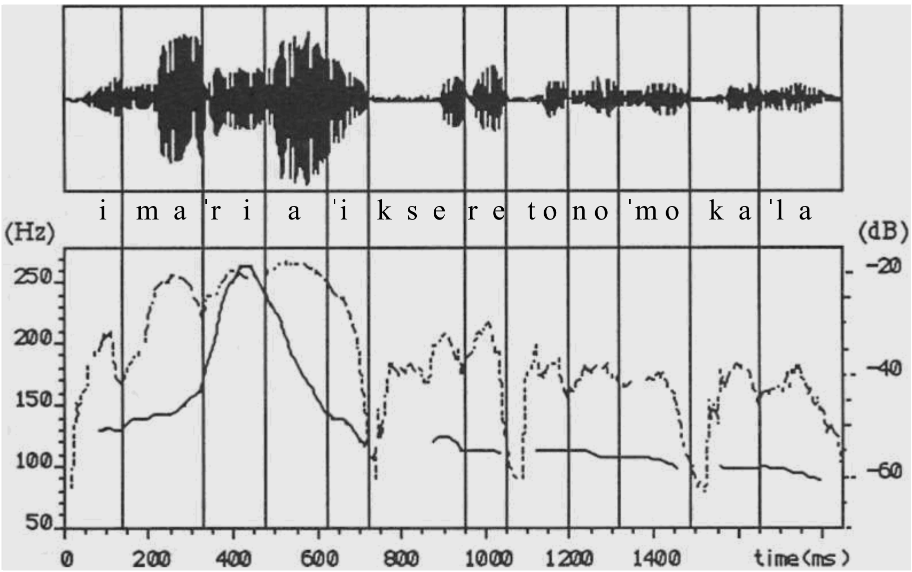


Figure 4.2. Waveform (over), F0 on the left axis (solid line) and intensity on the right axis (broken line) of the second member of the minimal pair [ˈnomo~noˈmo], in the carrier sentence [i maˈria ˈiksere to ___ kaˈla] in postfocal position.

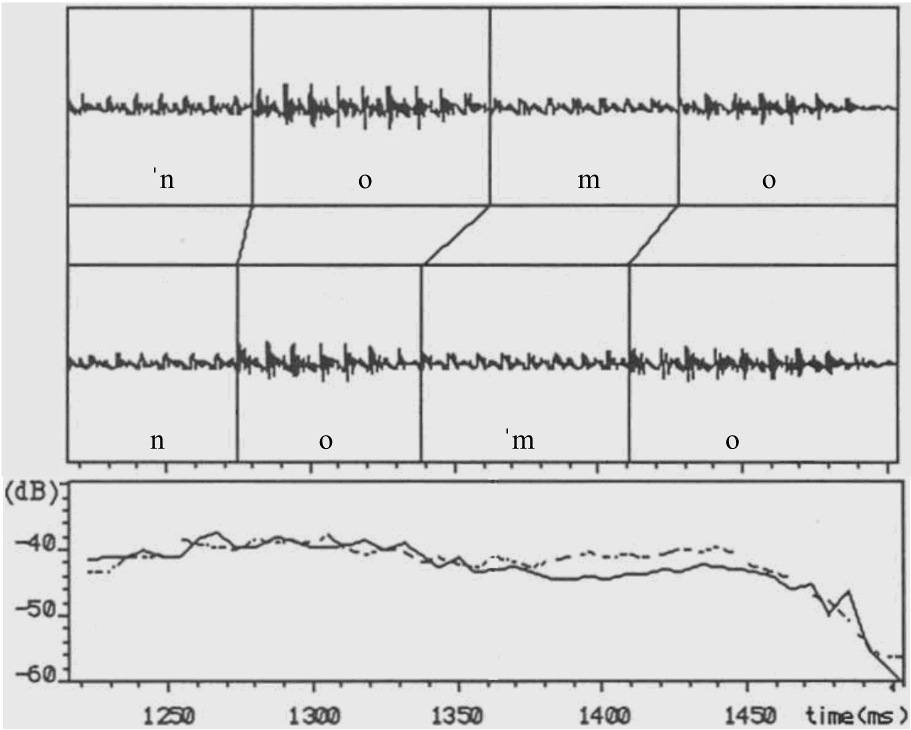


Figure 4.3. Waveform (over), and intensity contours (under) of the minimal pair [ˈnomo~noˈmo], in the carrier sentence [i **maˈria** ˈiksere to ˈnomo kaˈla] in postfocal position. The solid line shows the reference intensity contour of the first member of the pair and the broken line shows the target contour of the second one.

4.2.1.2 Results

Ten listeners' individual judgments of the ten stimuli in this experiment are shown in Table 4.2. St. 1, the [ˈnomo] sentence, and St. 2, the [noˈmo] sentence were correctly identified by 70% and 90%, respectively, i.e. listeners could perceive the word stress distinction after focus. St. 3, the [ˈnomo] word in [noˈmo] context was identified as [ˈnomo] by 70% and St. 4, the [noˈmo] word in [ˈnomo] context, was identified as [noˈmo] by 84%, i.e. there were no major context effects on the perception of postfocal word stress.

Table 4.1. Ten synthetic stimuli of the minimal pair [ˈnomo~noˈmo] in the carrier sentence [i **maˈria** ˈiksere to ___ kaˈla] in postfocal position.

		Reference		Variable	Target
St.	1	i **maˈria** ˈiksere to ˈnomo kaˈla			
St.	2	i **maˈria** ˈiksere to noˈmo kaˈla			
St.	3	ˈnomo	→	context substitution	noˈmo
St.	4	noˈmo	→	context substitution	ˈnomo
St.	5	ˈnomo	→	intensity manipulation	noˈmo
St.	6	ˈnomo	→	duration manipulation	noˈmo
St.	7	ˈnomo	→	intensity + duration	noˈmo
St.	8	noˈmo	→	intensity manipulation	ˈnomo
St.	9	noˈmo	→	duration manipulation	ˈnomo
St	10	noˈmo	→	intensity + duration	ˈnomo

Table 4.2. Ten listeners' responses (Sb) horizontally, for ten synthetic stimuli (St) vertically, of the minimal pair [ˈnomo~noˈmo] in the carrier sentence [i **maˈria** ˈiksere to ___ kaˈla] in postfocal position.

St \ Sb	01	02	03	04	05	06	07	08	09	10	ˈnomo %	01	02	03	04	05	06	07	08	09	10	noˈmo %
01	8	4	8	8	3	5	9	6	9	10	70	2	6	2	2	7	5	1	4	1	0	30
02	2	1	1	0	0	5	1	0	0	0	10	8	9	9	10	10	5	9	10	10	10	90
03	4	8	8	6	7	5	8	8	6	10	70	6	2	2	4	3	5	2	2	4	0	30
04	2	0	3	6	0	4	0	1	0	0	16	8	10	7	4	10	6	10	9	10	10	84
05	7	5	6	4	3	5	5	4	6	7	52	3	5	4	6	7	5	5	6	4	3	48
06	5	4	4	2	3	6	5	2	3	7	41	5	6	6	8	7	4	5	8	7	3	59
07	4	8	0	3	0	5	2	0	3	5	30	6	2	10	7	10	5	8	10	7	5	70
08	4	3	3	4	0	4	0	0	3	3	24	6	7	7	6	10	6	10	10	7	7	76
09	4	4	6	3	1	7	6	5	5	1	42	6	6	4	7	9	3	4	5	5	9	58
10	7	5	10	4	3	3	7	6	9	9	63	3	5	0	6	7	7	3	4	1	1	37

In stimuli 5, 6, and 7, we started from [ˈnomo] and the target word was [noˈmo]. We first changed intensity (St.5), which resulted in an identification change of 18% (from 70% to 52%); in a pairwise t-test with 9 degrees of freedom, this intensity change was significant ($t(9)=3.25$, $p<0.005$). Second, we changed duration (St.6) and obtained a 29% identification change (from 70% to 41%; $t(9)=3.71$, $p<0.005$). Finally, we changed both intensity and duration at the same time (St.7) and the result was a 40% identification change (from 70% to 30%; $t(9)=3.52$, $p<0.005$).

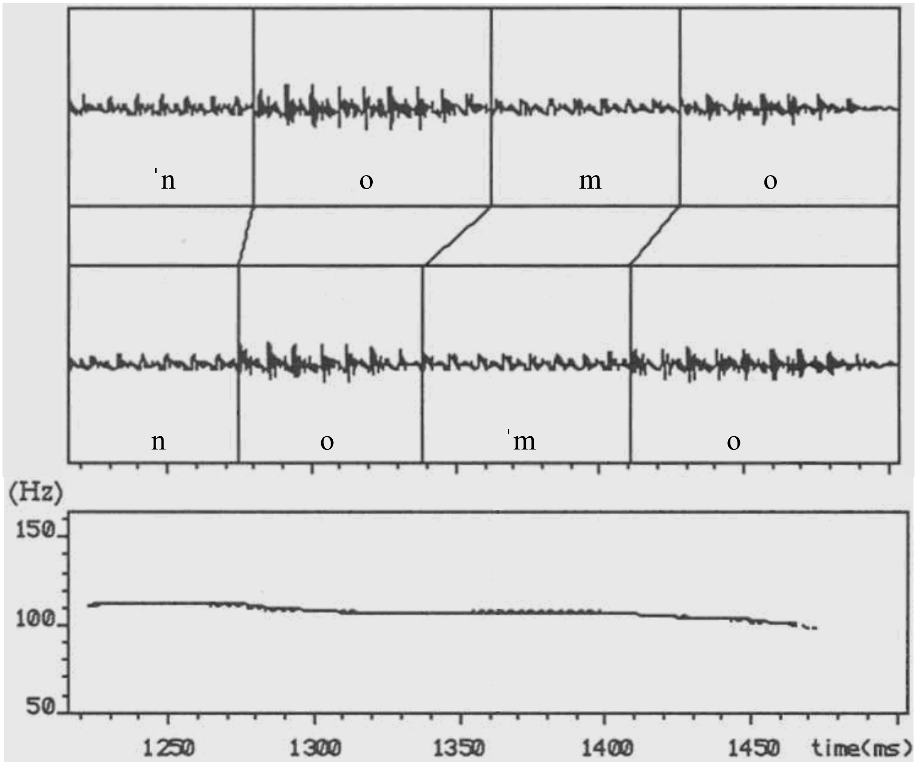


Figure 4.4. Waveform of the minimal pair [ˈnomo ~ noˈmo] (over) and F0 (under), in the carrier sentence [i **maˈria** ˈiksere to ___ kaˈla] in postfocal position. The solid line shows the reference F0 contour of the first member of the pair and the broken line shows the target F0 contour of the second member of the pair.

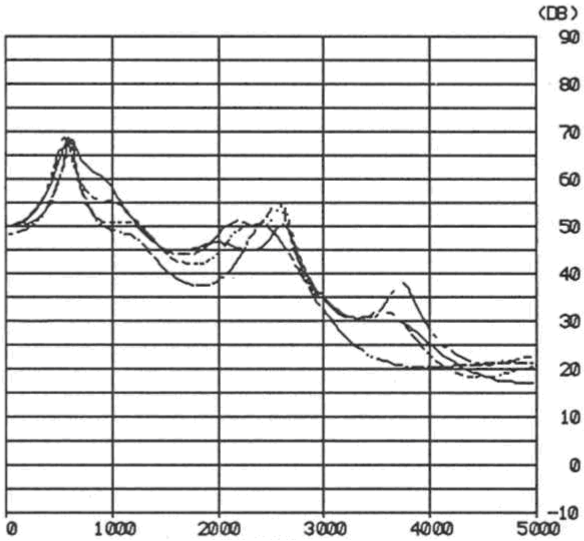


Figure 4.5. The formant structure of the four vowels of the minimal pair [ˈnomo~ nomó], in the carrier sentence [i **maˈria** ˈiksere to ___ kaˈla] in postfocal position.

In stimuli 8, 9, and 10, the reference word was [noˈmo] and the target word [ˈnomo]. The results of these three stimuli had the same structure as their mirror images, i.e. stimuli 5, 6, and 7. The identification changes of an intensity change (St.8) were 14% (from 90% to 76%; $t(9)=2.49$, $p<0.025$), a duration change (St.9) 32% (from 90% to 58%; $t(9)=6.00$, $p<0.0005$), and a combination of both intensity and duration changes (St.10) 53% (from 90% to 37%; $t(9)=4.92$, $p<0.0005$). Although the ten listeners did not react as a homogeneous group, there was no regular pattern of deviation, apart from listener 5 who was almost at a guessing level for both members of the pair.

An expansion of the F0 contour of [ˈnomo~noˈmo] showed no difference in F0 manifestation after focus, the two F0-contours being the same (Figure 4.4). The formant structure of the four vowels composing the minimal prosodic pair [ˈnomo~noˈmo] were much alike (Figure 4.5), so that the vowel quality could not have any effect to the perception of postfocal word stress.

4.2.2 Perceptual Cues of Prefocal Word Stress

4.2.2.1 Description of Stimuli

Ten synthetic stimuli were prepared for this experiment (Table 4.3). The listeners were asked to identify the investigated words as [ˈnomo] (law) or [noˈmo] (county), and mark with an acute accent the perceived stressed syllable of the grapheme {nomo} presented to them on an answer sheet of paper with ten rows, each row containing the ten stimuli in random order.

The first two stimuli were the human produced sentences [i maˈria ˈiksere to ˈnomo **kaˈla**] (Maria knew the law **well**) (Figure 4.6), and [i maˈria ˈiksere to noˈmo **kaˈla**] (Maria knew the county **well**) (Figure 4.7). The two sentences were answers to the question [pos ˈiksere i maˈria to ˈnomo ~ noˈmo] (how did Maria know the law ~ county?), i.e. the word [kaˈla] was in focus. The acoustic manifestation of the test sentences was quite regular, the minimal pair was produced as an independent tonal gesture with an F0 rise associated with the stressed syllables rather than as an assimilated or a compressed one; intensity as well as duration contributed on the syntagmatic and paradigmatic planes for the word stress distinction.

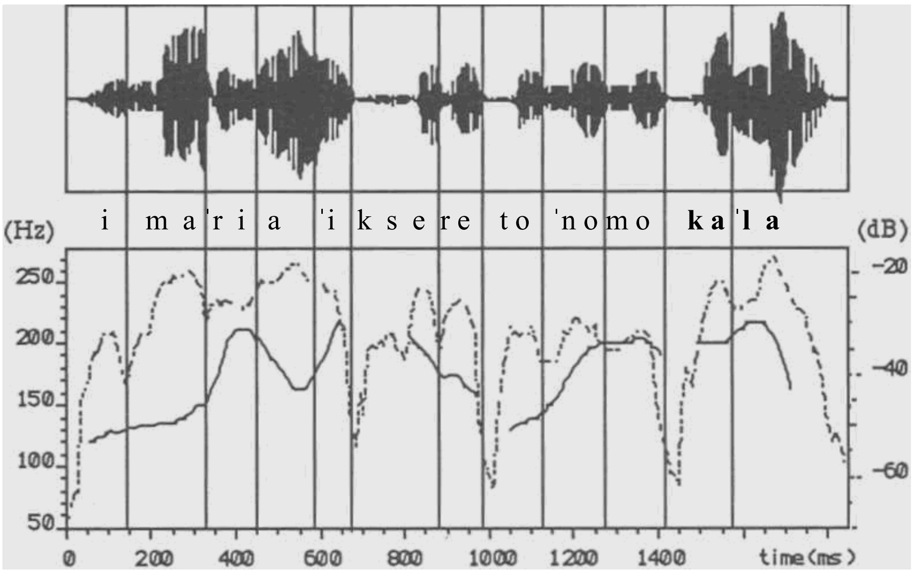


Figure 4.6.Waveform (over), F0 on the left axis (solid line) and intensity on the right axis (broken line) of the first member of the minimal pair [ˈnomo~noˈmo], in the carrier sentence [i maˈria ˈiksere to ___ **kaˈla**] in prefocal position.

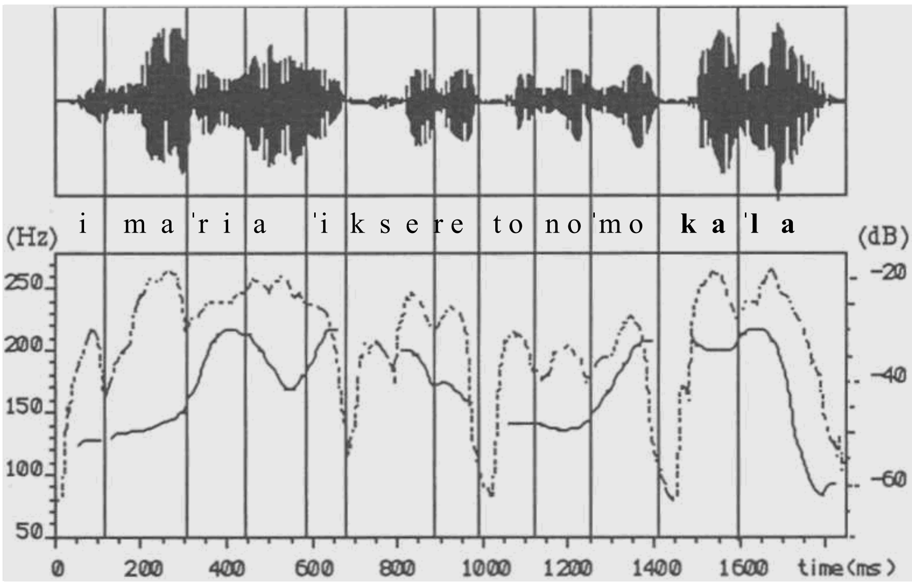


Figure 4.7. Waveform (over), F0 on the left axis (solid line) and intensity on the right axis (broken line) of the second member of the minimal pair [ˈnomo~noˈmo], in the carrier sentence [i maˈria ˈiksere to ___ **kaˈla**] in prefocal position.

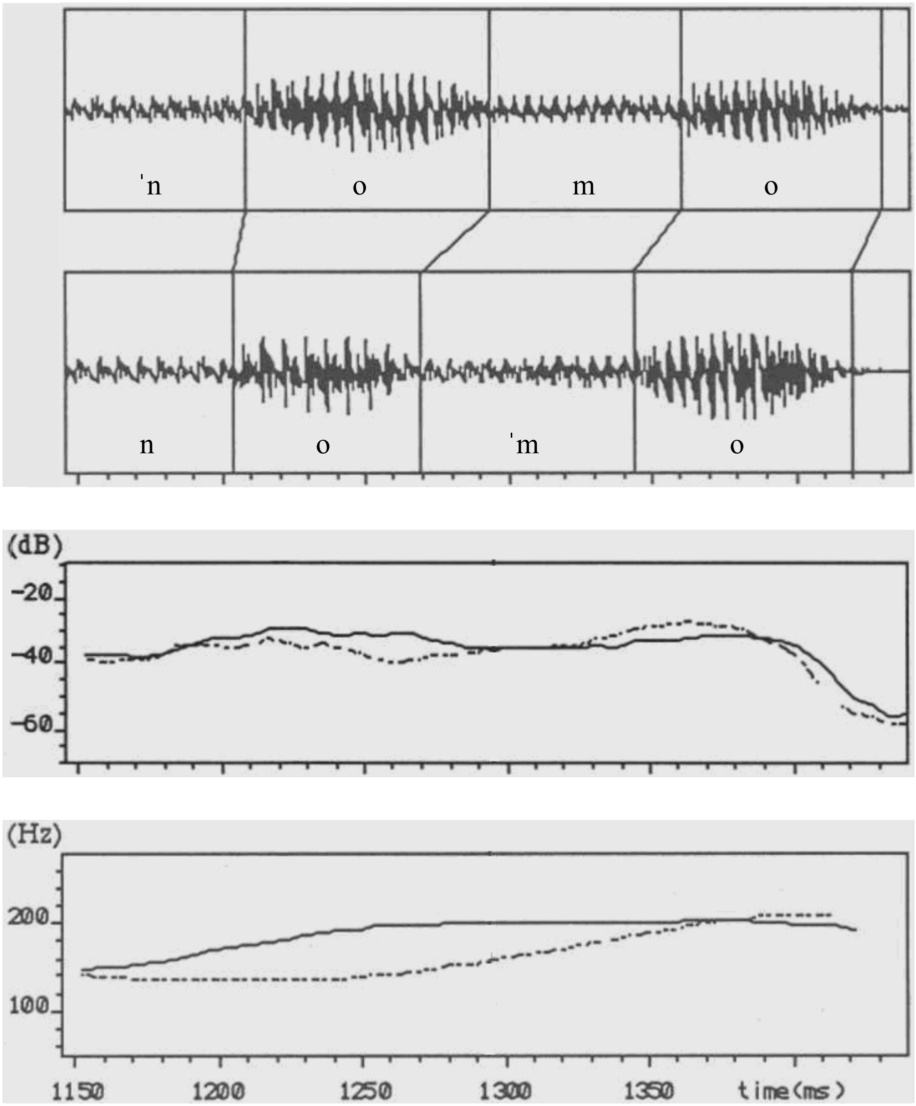


Figure 4.8. Waveform (over), intensity (middle) and F0 (under) of the minimal pair [ˈnomo~noˈmo], in the carrier sentence [i maˈria ˈksere to ˈnomo **kaˈla**], in prefocal position. The solid lines show intensity and fundamental frequency reference contours of the first member of the pair and the broken lines show the respective target contour of the second one.

In the following eight stimuli the acoustic signal of the minimal pair [ˈnomo~noˈmo] was expanded and the following manipulations were made: In St. 3 the intensity envelope of [ˈnomo] (ˈno=-28, mo=-31 dB) was changed to that of [noˈmo] (no=-32, ˈmo=-27 dB), (Figure 4.8). In St. 4 the four segments' duration of [ˈnomo] (n=57, o=83, m=64, o=70 ms) was changed to that of [noˈmo] (n=57, o=57, m=76, o=83 ms), (Figure 4.8). In St. 5 the F0 contour of [ˈnomo] (ˈno=140-200, mo=200-200 Hz) was changed to that of [noˈmo] (no=140-140, ˈmo=140-200 Hz), (Figure 4.8). In St. 6, all three acoustic parameters of intensity, duration and F0 of [ˈnomo] were changed to that of [noˈmo].

Stimuli 7, 8, 9, and 10 were mirror images of stimuli 3, 4, 5, and 6, respectively, i.e. this time changes were made from [noˈmo] and to [ˈnomo] changing only intensity (St. 7), only duration (St. 8), only F0 (St. 9), and finally all three parameters (I, D, F0) at the same time (St. 10). The above changes were restricted to the minimal pair [ˈnomo~noˈmo] leaving the context unaffected.

4.2.2.2 Results

The ten subjects' individual judgments of the ten stimuli in this experiment are shown in Table 4.4. St. 1, the [ˈnomo] sentence, and St. 2, the [noˈmo] sentence, were both identified correctly by 100%. Intensity change (St.3) as well as duration change (St.4) had no effect on the listeners' responses. F0 change (St.5) caused an overall identification change from [ˈnomo] to [noˈmo] as did an F0 change combined with intensity and duration changes (St.6). The responses of stimuli 7, 8, 9, and 10 had the same structure as their mirror images, stimuli 3, 4, 5, and 6. Intensity change (St.7) as well as duration change (St.8) had no effect; F0 change had an overall effect (St.9) as did F0 combined with intensity and duration changes (St. 10).

Table 4.3. Ten synthetic stimuli of the minimal pair [ˈnomo~noˈmo] in the carrier sentence [i maˈria ˈiksere to ___ **kaˈla**] in postfocal position.

St.	1	i maˈria ˈiksere to ˈnomo **kaˈla**			
St.	2	i maˈria ˈiksere to noˈmo **kaˈla**			
		Reference		Variable	Target
St.	3		ˈnomo →	Intensity (I)	noˈmo
St.	4		noˈmo →	Duration (D)	noˈmo
St.	5		ˈnomo →	Fundamental frequency (F0)	noˈmo
St.	6		noˈmo →	I+D+F0	noˈmo
St.	7		ˈnomo →	Intensity	ˈnomo
St.	8		noˈmo →	Duration	ˈnomo
St.	9		ˈnomo →	Fundamental frequency	ˈnomo
St.	10		noˈmo →	I+D+F0	ˈnomo

Table 4.4. Ten listeners' responses (Sb) horizontally, for ten synthetic stimuli (St) vertically, of the minimal pair [ˈnomo~noˈmo] in the carrier sentence [i maˈria ˈiksere to ___ **kaˈla**] in prefocal position.

Sb / St	01	02	03	04	05	06	07	08	09	10	ˈnomo %	01	02	03	04	05	06	07	08	09	10	noˈmo %
01	10	10	10	10	10	10	10	10	10	10	100	0	0	0	0	0	0	0	0	0	0	0
02	0	0	0	0	0	0	0	0	0	0	0	10	10	10	10	10	10	10	10	10	10	100
03	10	10	10	10	10	10	10	10	10	10	100	0	0	0	0	0	0	0	0	0	0	0
04	10	10	10	10	10	10	10	10	10	10	100	0	0	0	0	0	0	0	0	0	0	0
05	0	0	0	0	0	0	0	0	0	0	0	10	10	10	10	10	10	10	10	10	10	100
06	0	0	0	1	0	0	0	0	0	0	1	10	10	10	9	10	10	10	10	10	10	99
07	0	0	0	0	0	0	0	0	0	0	0	10	10	10	10	10	10	10	10	10	10	100
08	0	0	0	0	0	0	0	0	0	1	1	10	10	10	10	10	10	10	10	10	9	99
09	10	10	10	10	10	10	10	10	10	10	100	0	0	0	0	0	0	0	0	0	0	0
10	10	10	10	10	10	10	10	10	10	10	100	0	0	0	0	0	0	0	0	0	0	0

4.2.3 Prefocal Word Stress Perception: F0-range Dimension

4.2.3.1 Description of Stimuli

Fourteen synthetic stimuli were set up for this experiment (Table 4.5). The first group comprised seven stimuli with the word [ˈnomo] as the reference in the sentence [i maˈria ˈiksere to ˈnomo **kaˈla**] (Figure 4.9). In the first stimuli, F0 started rising from 140 Hz at the beginning of the stressed syllable to 200 Hz in the middle of the vowel of the stressed syllable and formed a plateau to the end of the word. In the following 6 stimuli, we kept the same starting point but we varied the plateau in equal steps of 10 Hz to the seventh stimuli which had a straight F0 contour.

The second group comprised seven stimuli as well, stimuli 8 to 14, and had the test word [noˈmo] as the reference in the sentence [i maˈria ˈiksere to noˈmo **kaˈla**] (Figure 4.10). The structure of the second group was the same as that of the first group, i.e. the F0 contour of the stressed syllable varied from 140 to 200 Hz in seven equal 10 Hz steps. The manipulations were restricted to the F0 contour of the minimal pair [ˈnomo ~ noˈmo], leaving intensity and duration as well as the test words' context unaffected.

4.2.3.2 Results

The ten listeners' individual judgments of the fourteen stimuli in this experiment are shown in Table 4.6. The results divide the stimuli into two groups; the first group, stimuli 1 to 7, was heard as [ˈnomo] and the second group, stimuli 8 to 14, was heard as [noˈmo]. The F0 manipulations had a minimal effect on listeners, St. 7 of the first group produced a 12% identification change whereas St. 11-14 of the second group caused less than a 10% change. It is clear that the F0 range does not affect considerably the listeners' responses for the two contrastive words which retain their original identity with intensity and duration.

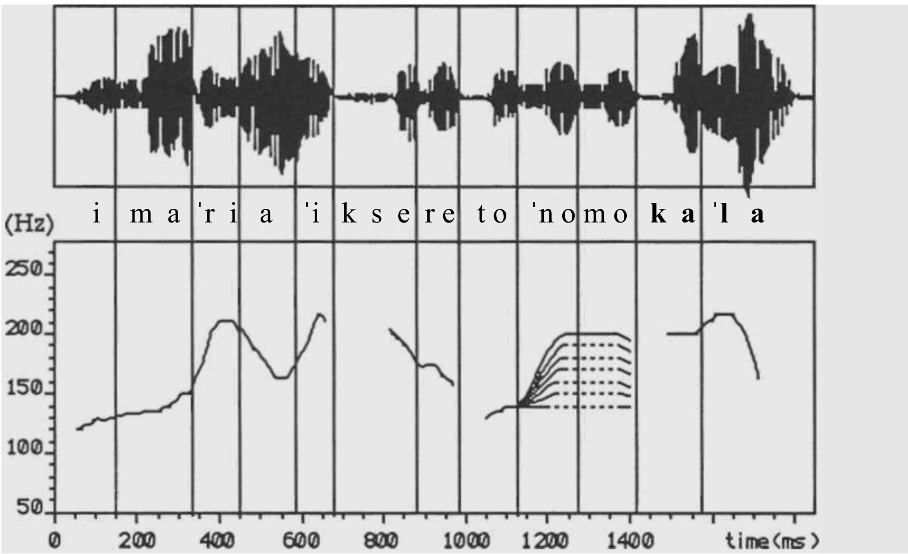


Figure 4.9. Waveform (over) and F0 (under) of the first member of the minimal pair [ˈnomo~noˈmo], in the carrier sentence [i maˈria ˈiksere to ___ **ka**ˈ**la**], in prefocal position; the solid line shows the reference contour and the broken lines show six synthetic stimuli in equal steps of 10 Hz.

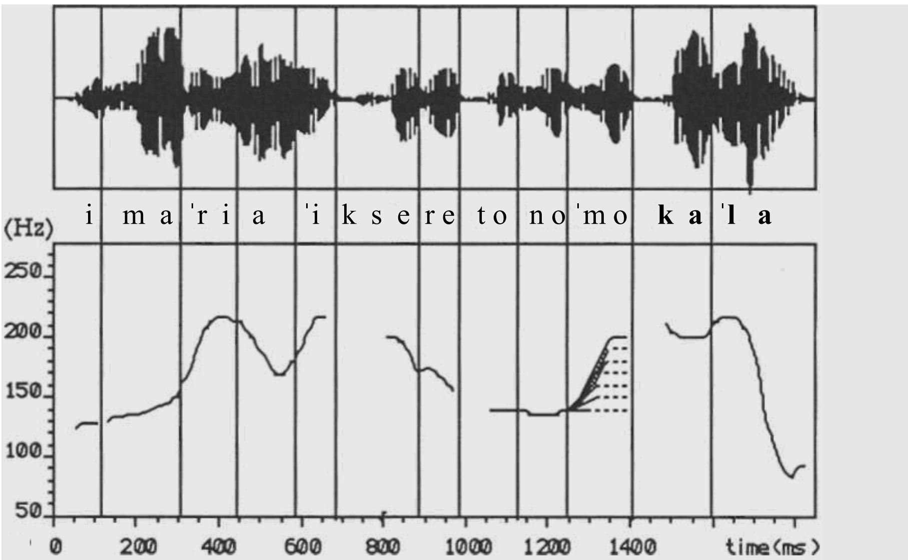


Figure 4.10. Waveform (over) and F0 (under) of the second member of the minimal pair [ˈnomo~noˈmo], in the carrier sentence [i maˈria ˈiksere to ___ **ka**ˈ**la**], in prefocal position; the solid line shows the reference contour and the broken lines show six synthetic stimuli in equal steps of 10 Hz.

Table 4.5. Fourteen synthetic stimuli of the minimal pair [ˈnomo ~ noˈmo] in the carrier sentence [i maˈria ˈiksere to ___ **kaˈla**] in prefocal position.

St. 1	(Hz)	200	ˈnomo	St. 8	(Hz)	200	noˈmo
St. 2		190	↓	St. 9		190	↓
St. 3		180		St. 10		180	
St. 4		170		St. 11		170	
St. 5		160		St. 12		160	
St. 6		150		St. 13		150	
St. 7		140		St. 14		140	

Table 4.6. Ten listeners' responses (Sb) for fourteen synthetic stimuli (St) of the minimal pair [ˈnomo ~ noˈmo] in the carrier sentence [i maˈria ˈiksere to ___ **kaˈla**] in prefocal position.

St \ Sb	01	02	03	04	05	06	07	08	09	10	ˈnomo %	01	02	03	04	05	06	07	08	09	10	noˈmo %
01	10	10	10	10	10	10	10	10	10	10	100	0	0	0	0	0	0	0	0	0	0	0
02	10	10	10	10	10	10	10	10	10	10	100	0	0	0	0	0	0	0	0	0	0	0
03	10	10	10	10	10	10	10	10	10	10	100	0	0	0	0	0	0	0	0	0	0	0
04	10	10	10	10	10	10	10	10	10	10	100	0	0	0	0	0	0	0	0	0	0	0
05	10	10	10	10	10	10	10	10	10	10	100	0	0	0	0	0	0	0	0	0	0	0
06	10	10	10	10	10	10	10	10	10	10	100	0	0	0	0	0	0	0	0	0	0	0
07	7	10	10	10	9	9	9	6	8	10	88	3	0	0	0	1	1	1	4	2	0	12
08	0	0	0	0	0	0	0	0	0	0	0	10	10	10	10	10	10	10	10	10	10	100
09	0	0	0	0	0	0	0	0	0	0	0	10	10	10	10	10	10	10	10	10	10	100
10	0	0	0	0	0	0	0	0	0	0	0	10	10	10	10	10	10	10	10	10	10	100
11	0	0	0	0	0	0	3	0	3	0	6	10	10	10	10	10	10	7	10	7	10	94
12	0	0	0	0	0	0	3	0	1	0	4	10	10	10	10	10	10	7	10	9	10	96
13	0	0	0	0	0	1	6	0	3	0	10	10	10	10	10	10	9	4	10	7	10	90
14	0	0	0	0	0	0	6	0	2	1	9	10	10	10	10	10	10	4	10	8	9	91

4.2.4 Prefocal Word Stress Perception: F0-timing Dimension

4.2.4.1 Description of Stimuli

Two groups, each with eight synthetic stimuli, were set up for this experiment (Table 4.7). The first group, stimuli 1-8, had the word [ˈnomo] as the reference in the sentence [i maˈria ˈiksere to ˈnomo **ka**ˈ**la**] (Maria knew the law **well**), an answer to the question [pos ˈiksere i maˈria to ˈnomo] (how did Maria know the law?). We started from [ˈnomo]’s F0 manifestation (reference) and making eight successive stimuli in equal steps we went forward to [noˈmo]’s F0 manifestation (target). All eight stimuli started at the same bottom of 140 Hz and ended at the same top of 200 Hz. The F0 rise of [noˈmo] was steeper than the F0 rise of [ˈnomo] and, consequently, the timing difference between the stimuli, to be proportional, was 3 frames (1 frame = 6.4 ms) for the bottom F0 and 2 frames for the top F0 (Figure 4.11).

The second group, stimuli 9-16, had the word [noˈmo] as the reference in the sentence [i maˈria ˈiksere to noˈmo **ka**ˈ**la**] (Maria knew the county **well**), an answer to the question [pos ˈiksere i maˈria to noˈmo] (how did Maria know the county?'). In the second group, we started from [noˈmo]’s F0 manifestation (reference) and went backwards in equal steps to [ˈnomo]’s F0 manifestation (target), (Figure 4.12). Both groups of stimuli were precisely the same, as far as F0 is concerned, so that each stimulus had its counterpart in the other group, i.e. St. 1 was identical with St. 16, St. 2 with 15, etc. (Table 4.1). The only difference between the two groups was that the first group, apart from F0, had [ˈnomo]’s acoustic manifestation whereas the second group had [noˈmo]’s one.

4.2.4.2 Results

The ten subjects’ individual judgments of the sixteen stimuli in this experiment are shown in Table 4.8. The first group of stimuli, stimuli 1-8, is divided into two categories; stimuli 1-5 are identified as [ˈnomo] with deviations less than 6%, St. 6 is ambiguous, and stimuli 7 and 8 are identified as [noˈmo]. The second group of stimuli, stimuli 9-16, is divided into two categories as well; stimuli 9-12 are identified as [noˈmo] with St. 12 deviating 6%, St. 13 is ambiguous (47%), and stimuli 14-16 are identified as [noˈmo], St. 14 deviating 6%. The per cent identification of the sixteen stimuli as [ˈnomo] or [noˈmo] is shown in Figure 4.13.

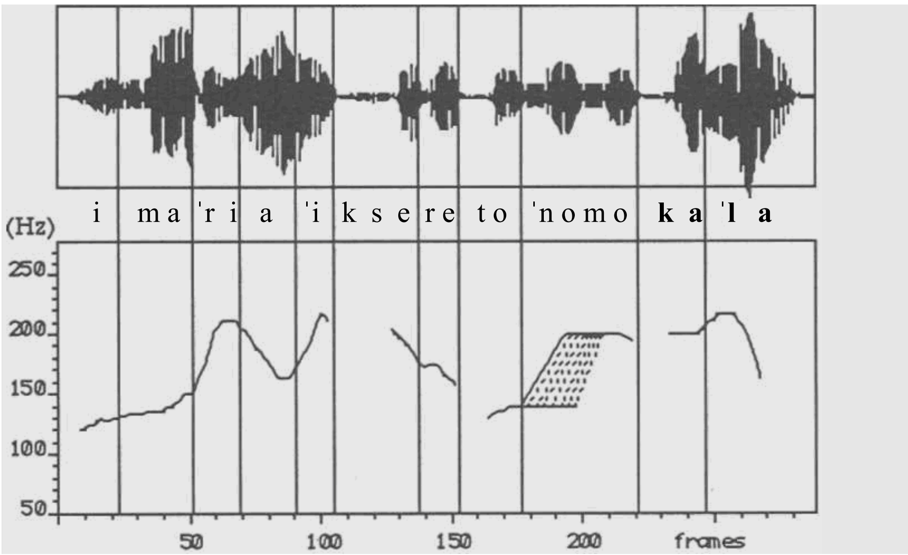


Figure 4.11. Waveform (over) and F0 (under) of the first member of the minimal pair [ˈnomo~noˈmo], in the carrier sentence [i maˈria ˈiksere to ___ **ka**ˈ**la**], in prefocal position; the solid line shows the reference F0 contour of the first member of the pair and the broken lines show six synthetic stimuli in 10 Hz steps.

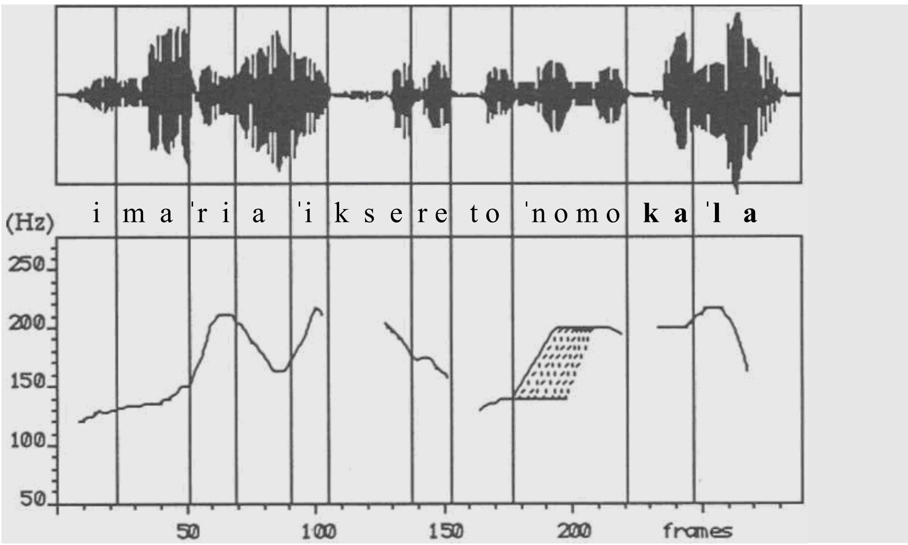


Figure 4.12. Waveform (over) and F0 (under) of the second member of the minimal pair [ˈnomo~noˈmo], in the carrier sentence [i maˈria ˈiksere to ___ **ka**ˈ**la**], in prefocal position; the solid line shows the reference F0 contour of the second member of the pair and the broken lines show six synthetic stimuli in 10 Hz steps.

Table 4.7. Sixteen synthetic stimuli of the minimal pair [ˈnomo~noˈmo] in the carrier sentence [i maˈria ˈiksere to ___ **kaˈla**], in prefocal position.

	Frames				Frames	
St. 1	180, 200	ˈnomo		St. 9	201, 214	noˈmo
St. 2	183, 202	↓		St. 10	198, 212	↓
St. 3	186, 204	↓		St. 11	195, 210	↓
St. 4	189, 206	↓		St. 12	192, 208	↓
St. 5	192, 208	↓		St. 13	189, 206	↓
St. 6	195, 210	↓		St. 14	186, 204	↓
St. 7	198, 212	↓		St. 15	183, 202	↓
St. 8	201, 214	noˈmo		St. 16	180, 200	ˈnomo

Table 4.8. Ten listeners' responses for sixteen synthetic stimuli of the minimal pair [ˈnomo~noˈmo] in the carrier sentence [i maˈria ˈiksere to ___ **kaˈla**] in prefocal position.

St \ Sb	01	02	03	04	05	06	07	08	09	10	ˈnomo %	01	02	03	04	05	06	07	08	09	10	noˈmo %
01	10	10	10	9	10	10	10	10	10	10		0	0	0	1	0	0	0	0	0	0	1
02	10	10	10	10	10	10	10	10	10	10	100	0	0	0	0	0	0	0	0	0	0	0
03	10	10	10	9	10	10	10	10	10	9	98	0	0	0	1	0	0	0	0	0	1	2
04	10	9	10	10	10	10	10	10	10	10	99	0	1	0	0	0	0	0	0	0	0	1
05	10	8	10	9	10	10	10	8	10	9	94	0	2	0	1	0	0	0	2	0	1	6
06	8	3	3	6	7	5	7	1	1	5	46	2	7	7	4	3	5	3	9	9	5	54
07	0	0	0	1	0	0	1	0	0	0	2	10	10	10	9	10	10	9	10	10	10	98
08	0	0	0	0	0	0	1	0	0	0	1	10	10	10	10	10	10	9	10	10	10	99
09	0	0	0	0	0	0	0	0	0	0	0	10	10	10	10	10	10	10	10	10	10	100
10	0	0	0	1	0	0	0	0	0	0	1	10	10	10	9	10	10	10	10	10	10	99
11	0	0	0	0	0	0	0	0	0	0	0	10	10	10	10	10	10	10	10	10	10	100
12	0	0	1	2	1	0	2	0	0	0	6	10	10	9	8	9	10	8	10	10	10	94
13	6	10	4	3	7	1	6	2	3	5	47	4	0	6	7	3	9	4	8	7	5	53
14	9	8	10	8	10	10	10	9	10	10	94	1	2	0	2	0	0	0	1	0	0	6
15	10	10	10	10	10	10	10	10	10	10	100	0	0	0	0	0	0	0	0	0	0	0
16	10	10	10	10	10	10	10	10	10	10	100	0	0	0	0	0	0	0	0	0	0	0

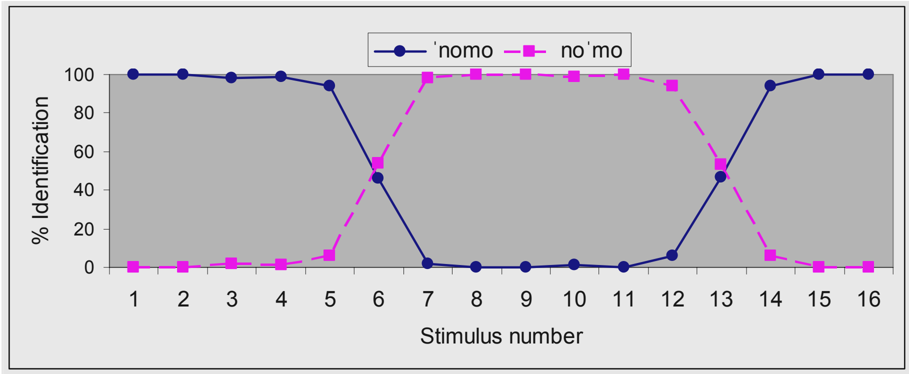


Figure 4.13. Percent identification of sixteen synthetic stimuli of the minimal pair [ˈnomo~noˈmo] in the carrier sentence [i maˈria ˈiksere to ___ **kaˈla**], in prefocal position.

4.2.5 Discussion

To recapitulate the results of the perceptual experiments on word stress distinction, duration had a greater effect on listeners than intensity in postfocal position whereas F0 outweighed both duration and intensity in prefocal position. On the other hand, a neutralization of F0 had a minor effect on listeners whereas the timing of the F0 rise associated with the word stress had an all-or-none effect in prefocal position.

In postfocal position (experiment I), neither F0 nor the vowel quality seems to contribute to the word stress distinction. This applies to both production and perception, the F0 contour and the formant structure of the investigated minimal pair being quite the same. As far as duration and intensity are concerned, although manipulating duration and intensity from [ˈnomo] to [noˈmo] and vice versa gave us different scores (Table 4.2), the structure of the listeners' responses was the same, i.e. a change of duration had a bigger effect than a change of intensity. Thus, duration was found to be a stronger perceptual cue than intensity. This is, however, rather weak evidence to use to attribute a higher rank in the hierarchy to duration since it cannot override the effects of intensity. It seems as if the perceptual system needs both duration and intensity, the energy integral (Beckman 1986) to make a clear decision and it is rather doubtful whether only one of these two parameters is strong enough to denote stress.

Although the listeners did not listen to the questions before the answers they could identify the intended answers correctly by 70% and 90% for each member of the pair, i.e. they could distinguish word stress after focus even in synthetic speech[7]. The exclusion of F0 for the postfocal word stress distinction is in accordance with the communicative power of the message. The test sentences were answers to the question [pça ˈiksere to ˈnomo ~ noˈmo kaˈla] (who knew the law ~ county well?). From a communicative point of view an answer like [i maˈria] (Maria (did)) or [i maˈria ton ˈiksere (Maria knew it) would be equally good. The rest of the sentence is already known to the questioner and consequently not so important to him.

The context had negligible influence on the listeners' decision, i.e. the investigated words kept their identity when they exchanged context. Our results are in disagreement with Huss (1978) who reports that the word stress distinction in English is maintained only when it is manifested by F0 whereas the distinction is neutralized in postfocal position where F0 is absent as a perceptual cue; moreover, in his experiments the contextual effect in terms of rhythmic structuring appeared to override the lexical stress distinction both in production and perception. In the Greek data, both production and perception respect the lexical distinction whereas a regular alternation of stressed ~ unstressed syllables would favour the first rather than the second member of the pair.

In prefocal position (experiment II), the hierarchic nature of the acoustic cues for word stress perception (Fry 1958, Hyman 1977, Berinstein 1979) has been validated. F0 has been found to be the all important prosodic factor to denote word stress whereas the relative contribution of the other parameters, intensity and duration, are totally overridden by the presence of F0. Thus, F0 appears as the main factor to determine prefocal word stress even if the other parameters are on the contrastive word, i.e. F0 totally overrides the conflicting cues of intensity and duration. Even if F0 was found to be the main prosodic parameter in this experiment, a classification of the other two, intensity and duration, remains unsettled. The

---

[7]In a preliminary perceptual experiment with the same material but a different recording, carried out in Lund, five Greek listeners could identify both members of the pair up to 90%.

predominance of F0 over duration and intensity combined is at odds with Beckman's (1986) results that duration and intensity as a single prosodic feature, the total amplitude, overrides F0 as a perceptual cue for nuclear stress in English.

It is already evident that the semantic/contextual structure of the language has a direct effect on the distribution of the acoustic parameters across the utterance and this is mainly witnessed in the F0 structure. The fact that F0 is hardly realized in postfocal position, that it may be assimilated in prefocal position but is realized with the largest variations in a constant way in focal position is an indication that the primary function of F0 is to contribute to the semantic weighting of certain speech elements against some others (Bruce 1985) rather than to contribute to word stress.

The variability of F0 across the utterance creates a major theoretical question as to whether F0 is really the primary acoustic cue, even for the prefocal word stress perception. If it is, why does the language not utilize the primary cue at the maximal extent at this position, i.e., why may the F0 contour appear in a rather compressed form? How can the listeners have a perfect stress identification with a neutralized F0? Why does the language not utilize F0 in postfocal position?

The most constant as well as the largest contribution of F0 across the Greek utterances is associated with sentence stress which is the prosodic category for focus bearing the most important information. This F0 variation is mainly realized on the stressed syllable rather than having as its domain the whole stress group, the lexical element or the focal part of the sentence the stressed syllable belongs to. Having established a direct relation between the largest F0 variation and the most important information we may argue that the stressed syllables may function as the "anchoring points" (Thorsen 1983) for the F0 changes whose primary function is to attribute to the constituent elements of the utterance the appropriate semantic weight.

> The presence of a suitably timed F0 event in connection with a stressed syllable is the accent that will mark the stressed syllable as more prominent and distinct from a stressed syllable without an accent. (Bruce 1983, p.224).

The dissociation of the F0 structure from the stress structure (Liberman and Pierrehumbert 1984) attributes the variation of F0 across the utterance to a semantic salience relation between its constituents. This eliminates the importance of F0 as the primary acoustic cue of stress, contrary to

Berinstein's hypothesis, and puts forward intension and duration, the energy integral (Fischer-Jørgensen 1984), as the primary acoustic cue for stress perception.

In experiment II, it was shown that F0 may have an all-or-none effect on the perception of word stress in prefocal position. If F0 were the only perceptual cue in this position, then St. 7 and St. 14 (experiment IΠ) with a straight F0-contour would both be perceived the same way, i.e. 50% as [ˈnomo] and 50% as [noˈmo]. Ideally, these figures would change along with the F0-changes either categorically or continuously, according to the mode of perception (Studdert-Kennedy et al. 1970)[8]. But this is obviously not the case. The contrastive words keep their original identity even with neutralized F0-contours.

The fact that a stressed syllable is not necessarily accompanied by an F0-change in production as well as the fact that the F0-contour may be neutralized without affecting stress perception is a strong indication that F0 is only indirectly associated with word stress. It rather seems that duration and intensity are quite constant word stress cues dividing the utterance into rhythmic units by an alternation of stressed and unstressed syllables (Bruce 1983, Strangert 1985). This is in agreement with stress detection results reported in Lieberman (1960), Lea (1977) and Beckman (1986) that the energy integral is a better predictor of stressed syllables than F0 is.

The rhythmic organization of the utterance is an essential condition for the F0-changes to take place, the stressed syllables coordinated with F0-changes which may be extended into the following unstressed syllables. Thus, F0 is attributed semantic and intonation properties (Pierrehumbert 1980) rather than rhythmic ones, the stressed ~ unstressed opposition.

The results of experiment IV have validated the all-or-none effect of F0 moving across the minimal pair [ˈnomo~noˈmo] in prefocal position. The timing of the F0-rise causes a complete identification change, totally

---

[8]Categorical perception according to Studdert-Kennedy et al. (1970) must meet the following three criteria: (1) sharp categorization, (2) peak discrimination at the category boundary, and (3) troughs in discrimination within categories. Here categorical perception is used in a rather broad sense meaning abrupt identification change, nearly corresponding to the first of the above mentioned criteria.

overriding the conflicting cues of duration and intensity. This may seem contradictory to our argument in experiment ΠΙ that it is duration and intensity which function as perceptual cues for word stress. However, this paradox may be resolved if we consider F0 as a higher order acoustic parameter whose function is to contribute to the semantic weighting of a certain part of the utterance in relation to the others (Bruce 1985), rather than as an acoustic cue for word stress (Berinstein 1979).

In a speaker ~ hearer relation, the native speaker of Greek “knows” that an F0-rise is temporally coordinated with a stressed syllable. The direct effect of the F0-rise displacement is a stress identification change in accordance with the alignment of the F0-rise with the corresponding syllable.

The F0-rise moving across the minimal pair [ˈnomo~noˈmo] is perceived categorically. For both groups of stimuli it is only one stimuli which is ambiguous, all other stimuli belonging either to one category or the other, i.e. we have an identification change within 6 frames (38,4 ms) for the bottom and 4 frames (25,6 ms) for the top F0. Categorical perception has been reported for Swedish between acute and grave accent (Bruce 1977) as well as for English nuclear tones (Lindsay and Ainsworth 1985).

However, it is important to notice that the critical point of identification change is not the same for both groups. For instance, St. 6 which belongs to the first group of stimuli having [ˈnomo] as the reference is ambiguous; on the other hand, St. 11 which has the same F0 manifestation as St. 6 is identified as [noˈmo]. The same relation holds for stimuli 13 and 4; although they have the same F0 manifestation, St. 13 is ambiguous but St. 4, which has [ˈnomo] as the reference, is identified as [ˈnomo].

We can explain the above phenomenon if we put it in relation to the argument of experiment ΠΙ, i.e. F0 is a higher order acoustic parameter with an all-or-none effect on stress perception. When we start from [ˈnomo] and go to [noˈmo] moving F0 horizontally, we have to come clear over the mid-point so that F0 neutralizes the effect of duration and intensity on [ˈnomo]. Thus, stimuli 4 and 5, although they are at the mid-point of the F0 variations, are both identified as [ˈnomo] since they have [ˈnomo] as the reference. On the other hand, stimuli 12 and 13 which are identical to stimuli 4 and 5 respectively, as far as F0 is concerned, but have [noˈmo] as the reference, have another identification; St. 12 is identified as [noˈmo] whereas St. 13 is ambiguous. This finding reflects the influence that duration and intensity have on prefocal word stress perception.

4.3 Perception of Enclitic Stress

4.3.0 Introduction

The final two perceptual experiments (V & VI) were concerned with enclitic stress perception. In experiment V two syntactic structures were set up. The first one was a proclitic structure with a word stress and the second one an enclitic structure with a word and an enclitic stress associated with it. The first question to answer was if enclitic stress is a phonetic reality, in other words if listeners perceive enclitic stress and associate it with its corresponding syntactic structure, in the present experiment even in synthetic speech. The second question was to find out which one of the three acoustic parameters, i.e. intensity, duration or F0, is the strongest perceptual cue and if we can establish a hierarchy among the acoustic parameters; moreover, since the enclitic stressed syllable in the investigated speech material is manifested by an F0 rise and an F0 fall, as far as F0 is concerned, the perceptual contribution of the F0 fall was investigated as well. Last, the third question to answer was whether only one of the acoustic parameters is able to convey enclitic stress or if a combination of more acoustic parameters is an absolute necessity.

In experiment VI, led by the results of experiment V, we investigated the perceptual relevance of F0 for the enclitic stress distinction and its effect on the proclitic ~ enclitic opposition, testing synthetic stimuli in successive steps from the proclitic manifestation of the F0-contour to the enclitic one and vice versa.

4.3.1 Perceptual Cues of Enclitic Stress

4.3.1.1 Description of Stimuli

Twelve synthetic stimuli were set up for this experiment (Table 4.9). The first two stimuli were the sentences [to ˈonoma mu ˈitane ɣnoˈsto] (the name was familiar to me) (Figure 4.14), and [to ˈonoˈma mu ˈitane ɣnoˈsto] (my name was familiar (to them)), (Figure 4.15).

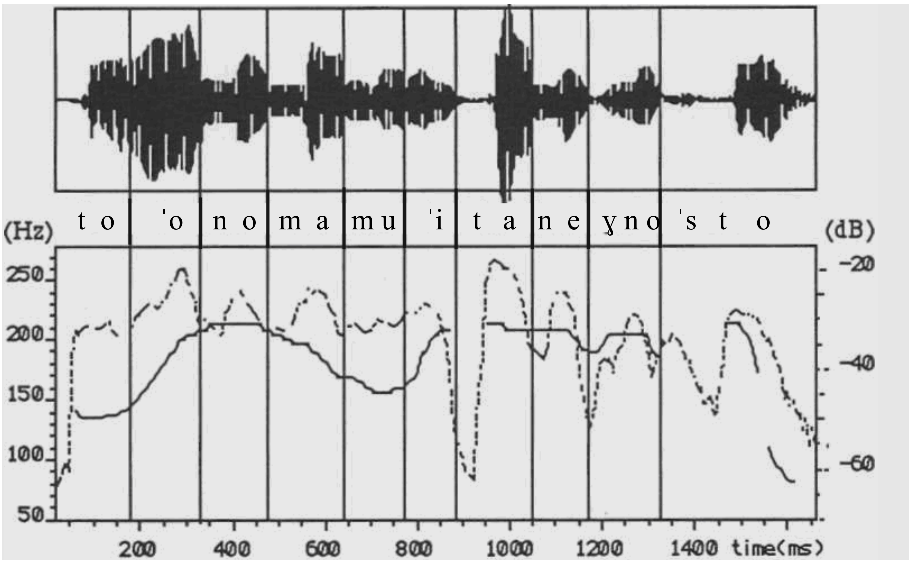


Figure 4.14. Waveform (over), F0 on the left axis (solid line) and intensity on the right axis (broken line) of the first member of the minimal pair [ˈonoma~ˈonoˈma], in the carrier sentence [to ___ mu ˈitane ɣnoˈsto].

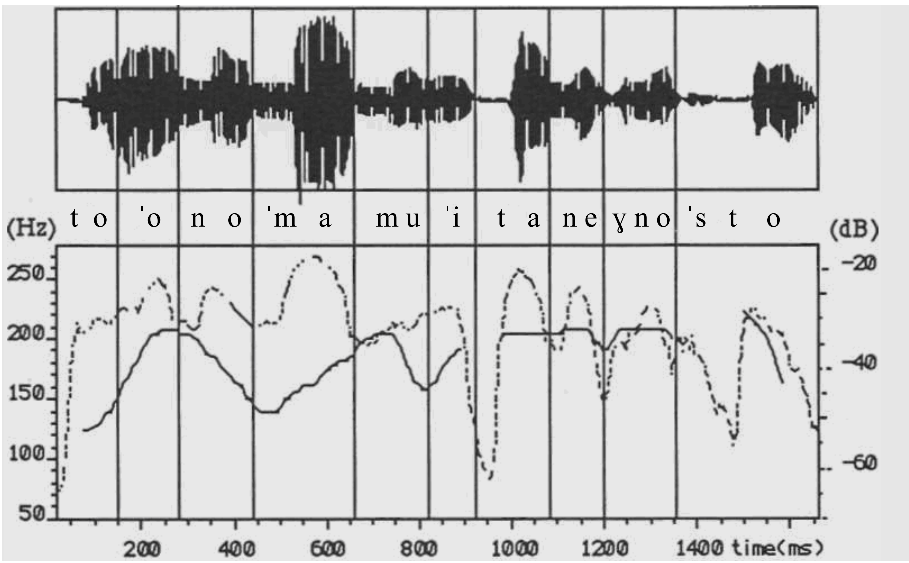


Figure 4.15. Waveform (over), F0 on the left axis (solid line) and intensity on the right axis (broken line) of the second member of the minimal pair [ˈonoma~ˈonoˈma], in the carrier sentence [to ___ mu ˈitane ɣnoˈsto].

In the first sentence the word [ˈonoma] carries a word stress and the clitic [mu] is a personal pronoun belonging syntactically to the following verb phrase, i.e. [mu] functions as a proclitic element. In the second sentence, the word [ˈonoˈma] carries two stresses, a word stress and an enclitic stress, and the clitic [mu] is a possessive pronoun belonging to the noun phrase, i.e. the clitic [mu] functions as an enclitic element having the syntactic boundary after it. The two sentences were pronounced neutrally, i.e. the speaker had no contextual information. The minimal pair under investigation [ˈonoma~ˈonoˈma] appeared with a rather regular acoustic manifestation, one stress group with an independent tonal gesture for the proclitic structure and two stress groups with two independent tonal gestures for the enclitic structure; intensity and duration made their regular contribution to the word stress of the proclitic structure as well as to the word and enclitic stress of the enclitic structure respectively (see part two).

In the rest of the stimuli we made the following manipulations by expanding the acoustic signal of the minimal pair [ˈonoma~ˈonoˈma]: In St. 3 we changed the intensity envelope of [ˈonoma (ˈo=-19, no=-24, ma=-24 dB) to that of [ˈonoˈma] (ˈo=-22, no=-24, ma=-17 dB), (Figure 4.16). In St. 4 we changed the five segments' duration of [ˈonoma] (ˈo=166, n=70, o=77, m=83, a=77 ms) to that of [ˈonoˈma] (ˈo=121, n=70, o=83, m=96, a=128 ms), (Figure 4.16). The F0 of the minimal pair [ˈonoma~ˈonoˈma] (Figure 4.16) was manipulated in two steps. In the first step, St. 5, we changed the F0 rise and F0 fall of [ˈonoma] to that of [ˈonoˈma], to see the effects of the final F0 fall on the phrase boundary and consequently on the enclitic stress; in the second step, St. 6, in addition to the initial F0 rise and the final F0 fall we added a fall-rise associated with the enclitic stress, i.e. St. 6 included St. 5 (Figure 4.17). Thus, into two steps, stimuli 5 and 6, we turned the one-stressed (one F0 top) [ˈonoma] into two stressed [ˈonoˈma] (two F0 tops). The F0 variations were from 140 to 200 Hz (Figure 4.17). St. 7 was a combination of stimuli 3, 4, 5, and 6, i.e., we changed intensity, duration, and F0 of [ˈonoma] to that of [ˈonoˈma].

Stimuli 8, 9, 10, 11, and 12 were mirror images of stimuli 3, 4, 5, 6, and 7, i.e., we started from [ˈonoˈma] and went to [ˈonoma] changing intensity (St.8), duration (St.9), F0 step 1 (St. 10), F0 step 2 (St. 11), and finally intensity, duration and F0 (St. 12).

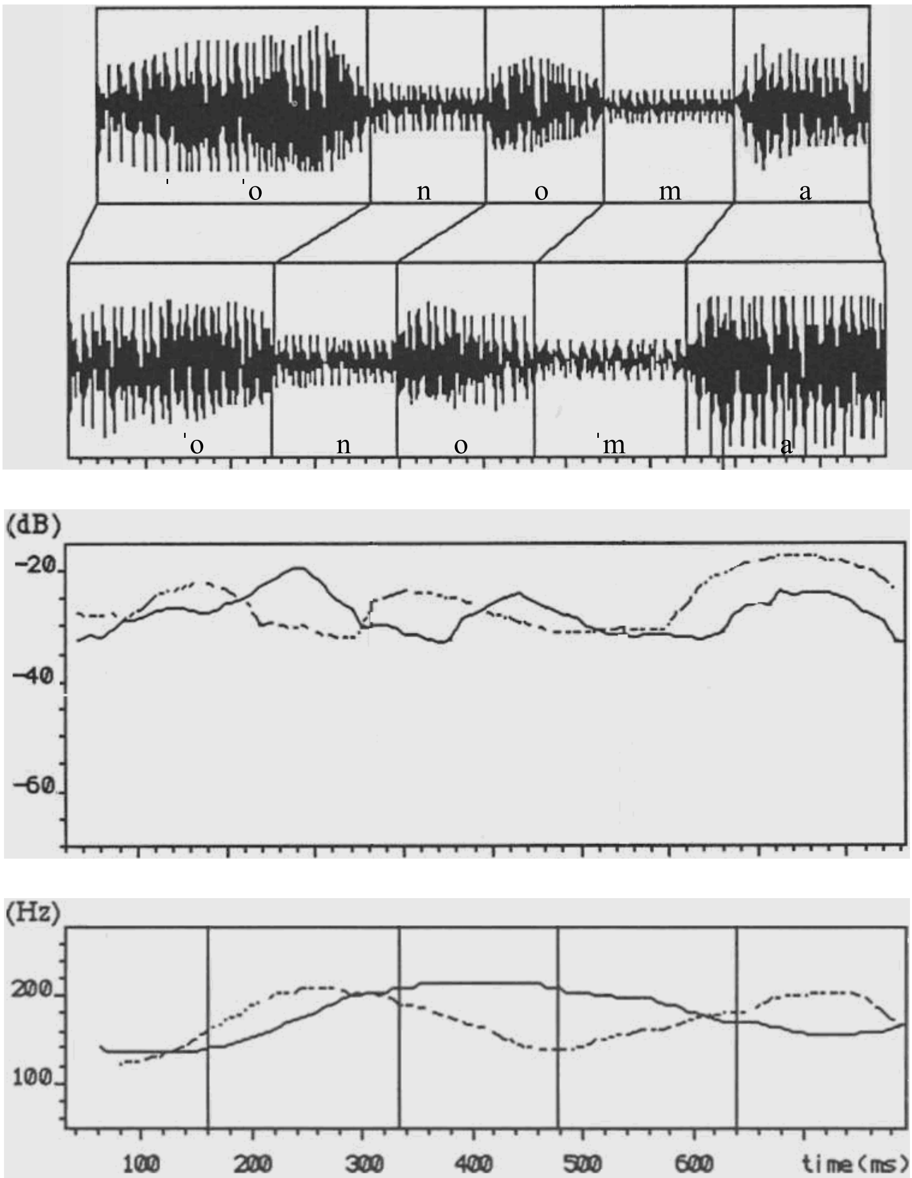


Figure 4.16. Waveform (over), intensity (middle) and F0 (under) of the minimal pair [ˈonoma~ˈonoˈma], in the carrier sentence [to ___ mu ˈitane ɣnoˈsto]. The solid lines show intensity and fundamental frequency reference contours of the first member of the pair and the broken lines show the respective target contour of the second one.

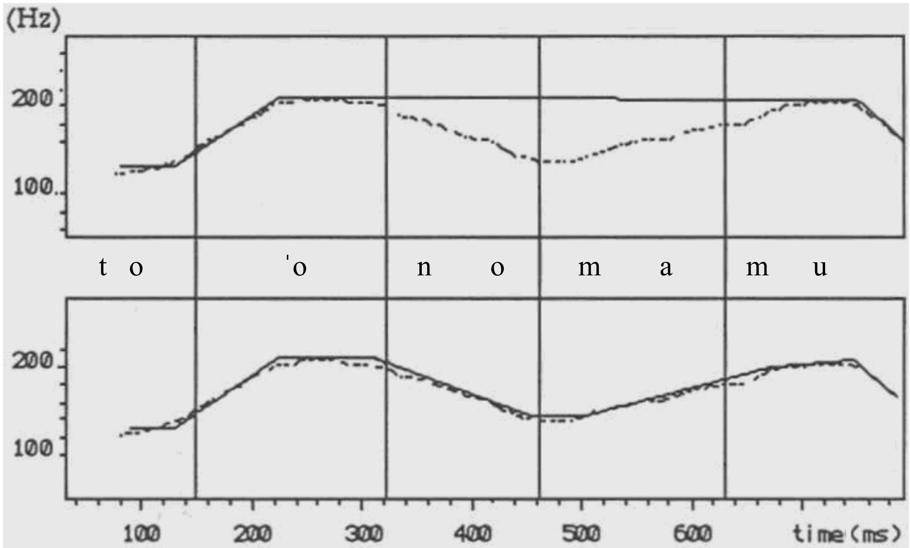


Figure 4.17. F0 contours of the minimal pair [ˈonoma~ˈonoˈma], in the carrier sentence [to ___ ˈitane ɣnoˈsto]. The solid line shows the 1-step (over) and the 2-step (under) F0 synthetic manipulations from the first to the second member of the pair (broken line).

4.3.1.2 Results

Ten subjects' individual identifications of the twelve stimuli in this experiment are shown in Table 4.10.

St. 1, the proclitic structure [to ˈonoma mu ˈitane ɣnoˈsto], and St. 2, the enclitic structure [to ˈonoˈma mu ˈitane ɣnoˈsto], were correctly identified by 97% and 100% respectively. In stimuli 3, 4, 5, 6, and 7 we started from [ˈonoma], the proclitic structure, and the target was [ˈonoˈma], the enclitic structure. In St. 3, the intensity change had no effect on the listeners' responses (from 97% to 98%). In St. 4, the duration change produced an identification change of 24% (from 97% to 73%; $t(9)=4.43$, $p<0.005$). In St. 5, the F0 step l manipulation, initial F0 rise and final F0 fall, had no effect (from 97% to 96%). In St. 6 the F0 step 2 manipulation, initial F0 rise, F0 fall-rise associated with the enclitic stress, and final F0 fall, had an effect of 32% (from 97% to 65%; $t(9)=5.23$, $p<0.0005$). St. 7, a combination of intensity, duration, and F0 manipulations, produced a complete identification change (from 97% to 2%).

Table 4.9. Twelve synthetic stimuli of the minimal pair [ˈonoma~ˈonoˈma] in the carrier sentence [to ___ mu ˈitane ɣnoˈsto].

St.		Reference		Variable	Target
St.	01	to ˈonoma mu ˈitane ɣnoˈsto			
St.	02	to ˈonoˈma mu ˈitane ɣnoˈsto			
St.	03	ˈonoma	→	Intensity (I)	ˈonoˈma
St.	04	ˈonoma	→	Duration (D)	ˈonoˈma
St.	05	ˈonoma	→	F0-step 1	ˈonoˈma
St.	06	ˈonoma	→	F0-step 2	ˈonoˈma
St.	07	ˈonoma	→	I+D+F0	ˈonoˈma
St.	08	ˈonoˈma	→	Intensity	ˈonoma
St.	09	ˈonoˈma	→	Duration	ˈonoma
St.	10	ˈonoˈma	→	F0-step 1	ˈonoma
St.	11	ˈonoˈma		F0-step 2	ˈonoma
St.	12	ˈonoˈma		I+D+F0	ˈonoma

Table 4.10. Ten listeners' responses (Sb) for fourteen synthetic stimuli (St) of the minimal pair [ˈonoma~ˈonoˈma] in the carrier sentence [to ___ mu ˈitane ɣnoˈsto].

St \ Sb	01	02	03	04	05	06	07	08	09	10	ˈonoma %	01	02	03	04	05	06	07	08	09	10	ˈonoˈma %
01	10	10	10	10	10	9	8	10	10	10	99	0	0	0	0	0	0	1	0	0	0	1
02	0	0	0	0	0	0	0	0	0	0	91	1	1	0	3	0	2	2	0	0	0	9
03	10	10	10	10	10	10	10	9	10	9	79	0	4	2	3	0	4	2	3	2	1	21
04	5	8	9	8	8	10	4	7	8	6	80	0	1	1	1	0	3	5	7	1	1	20
05	10	10	10	10	10	9	9	10	9	9	79	0	1	2	4	0	4	1	5	3	1	21
06	8	8	7	4	8	6	8	7	4	5	78	1	3	1	3	0	3	2	8	1	0	22
07	0	0	0	0	0	0	1	1	0	0	71	0	3	1	2	0	2	1	7	3	10	29
08	0	0	0	0	0	0	1	0	0	0	2	10	9	10	10	10	10	9	10	9	10	98
09	1	0	0	0	0	0	1	1	0	2	1	10	9	10	10	10	10	10	10	10	10	99
10	0	0	0	0	0	1	1	1	0	0	2	10	9	9	10	10	10	10	10	10	10	98
11	2	6	5	5	6	6	4	4	2	1	3	10	9	10	10	10	10	8	0	0	0	97
12	10	10	10	10	10	10	10	9	10	8	5	10	9	9	10	10	10	9	10	8	10	95

Stimuli 8, 9, 10, 11, 12 were mirror images of stimuli 3, 4, 5, 6, 7, i.e., starting from [ˈonoˈma] to the target [ˈonoma]. Intensity change had no effect (St. 8, from 100% to 99%), duration change had a negligible effect (St. 9, from 100% to 95%) as well as F0 step 1 (St. 10, from 100% to 97%). F0 step 2 change caused a 41% identification change (St. 11, from 100% to 59%; $t(9)=6.99$, $p<0.0005$). Finally, a combination of intensity, duration, and F0 caused a complete identification change (St. 12, from 100% to 3%).

4.3.2 Enclitic Stress Perception: F0-dimension

4.3.2.1 Description of Stimuli

Fourteen synthetic stimuli were set up for this experiment (Table 4.11). The seven first stimuli had [ˈonoma] as the reference in the proclitic structure [to ˈonoma mu ˈitane ɣnoˈsto] (the name was familiar to me); we started from St. 1, the F0 manifestation of [ˈonoma], and by changing the F0 plateau into a fall-rise through six successive stimuli (F0 was lowered in equal steps of 10 Hz from 200 to 140 Hz) we came to [ˈonoˈma], the enclitic structure of [to ˈonoˈma mu ˈitane ɣnoˈsto] (my name was familiar), (Figure 4.18). The remaining seven stimuli had the enclitic structure as the reference; we started with St. 8, the F0 manifestation of the enclitic structure, and by raising F0 in six equal steps of 10 Hz, from 140 to 200 Hz, we came to the proclitic structure (Figure 4.19). The F0 manipulations were restricted to the domain of the contrastive phrases; the other acoustic parameters were unaffected.

4.3.2.2 Results

Ten listeners' individual judgments of the fourteen synthetic stimuli in this experiment are shown in Table 4.12.

The results divide the stimuli into two groups. The first group, stimuli 1 to 7, was identified as the proclitic sentence [to ˈonoma mu ˈitane ɣnoˈsto], i.e. the reference utterance. Although the F0 manipulations had an effect of 28% on the listeners' responses from St. 1 to St. 7, it is not clear whether we have continuous or abrupt identification changes. The second group, stimuli 8 to 14, was identified as the enclitic sentence [to ˈonoˈma mu ˈitane ɣnoˈsto], i.e. the reference utterance. The F0 manipulations had a minor effect on stimuli 8 to 13 and an 18% effect on St. 14. If we compare the scores of the second group with those of the first one we see that they do not correspond; moreover, the results do not show any clear tendency towards perceptual grouping of the stimuli between the two groups.

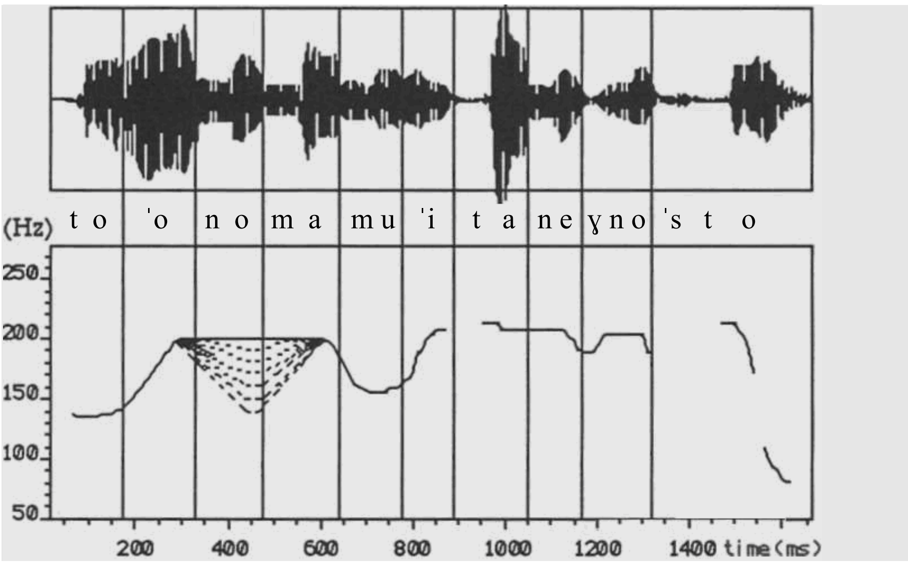


Figure 4.18. Waveform (over) and F0 (under) of the first member of the pair [ˈonoma~ˈonoˈma], in the carrier sentence [to ___ mu ˈitane ɣnoˈsto]. The solid line shows the reference F0 contour of the first member of the pair and the broken lines show six synthetic stimuli to the target F0 contour of the second one.

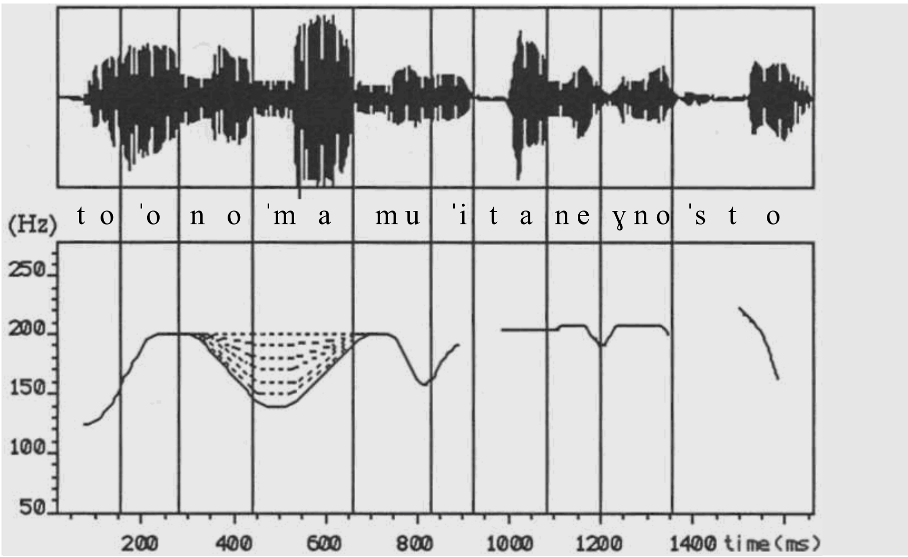


Figure 4.19. Waveform (over) and F0 (under) of the first member of the pair [ˈonoma~ˈonoˈma], in the carrier sentence [to ___ mu ˈitane ɣnoˈsto]. The solid line shows the reference F0 contour of the first member of the pair and the broken lines show six synthetic stimuli to the target F0 contour of the second one.

Table 4.11. Fourteen synthetic stimuli of the minimal pair [ˈonoma~ˈonoˈma] in the carrier sentence [to ___ mɯ ˈitane ɣnoˈsto].

St. 1	(Hz)	200	ˈonoma	St. 8	(Hz)	200	ˈonoˈma
St. 2		190	↓	St. 9		190	↓
St. 3		180	↓	St. 10		180	↓
St. 4		170	↓	St. 11		170	↓
St. 5		160	↓	St. 12		160	↓
St. 6		150	↓	St. 13		150	↓
St. 7		140	ˈonoˈma	St. 14		140	ˈonoma

Table 4.12. Ten listeners' responses (Sb) for fourteen synthetic stimuli (St) of the minimal pair [ˈonoma~ˈonoˈma] in the carrier sentence [to ___ mu ˈitane ɣnoˈsto].

St \ Sb	01	02	03	04	05	06	07	08	09	10	ˈonoma %	01	02	03	04	05	06	07	08	09	10	ˈonoˈma %
01	10	10	10	10	10	10	9	10	10	10	99	0	0	0	0	0	0	1	0	0	0	1
02	9	9	10	7	10	8	8	10	10	10	91	1	1	0	3	0	2	2	0	0	0	9
03	10	6	8	7	10	6	8	7	8	9	79	0	4	2	3	0	4	2	3	2	1	21
04	10	9	9	9	10	7	5	3	9	9	80	0	1	1	1	0	3	5	7	1	1	20
05	10	9	8	6	10	6	9	5	7	9	79	0	1	2	4	0	4	1	5	3	1	21
06	9	7	9	7	10	7	8	2	9	10	78	1	3	1	3	0	3	2	8	1	0	22
07	10	7	9	8	10	8	9	3	7	0	71	0	3	1	2	0	2	1	7	3	10	29
08	0	1	0	0	0	0	0	0	1	0	2	10	9	10	10	10	10	9	10	9	10	98
09	0	1	0	0	0	0	0	0	0	0	1	10	9	10	10	10	10	10	10	10	10	99
10	0	1	1	0	0	0	0	0	0	0	2	10	9	9	10	10	10	10	10	10	10	98
11	0	1	0	0	0	0	2	0	0	0	3	10	9	10	10	10	10	8	0	0	0	97
12	0	1	1	0	0	0	1	0	2	0	5	10	9	9	10	10	10	9	10	8	10	95
13	0	0	1	1	0	1	0	0	1	0	4	10	10	9	9	10	9	10	10	9	10	96
14	1	2	2	3	5	0	3	0	3	1	20	9	8	8	7	5	10	7	10	7	9	80

4.3.3 Discussion

The all-or-none effect of F0 on word stress does not hold for enclitic stress. It rather seems that this higher level acoustic parameter cannot override the lower level acoustic parameters of duration and intensity for enclitic stress. On the other hand, we have observed in part two of this study that enclitic stress affects the focal contribution of F0 within the enclitic structure. This is in accordance with the rhythmosyntactic nature of enclitic stress which is provided by syntax on rhythmic grounds and not by the lexicon, i.e. the relation of rhythm to tonal structure has different dimensions for word and enclitic stress.

In the case of the minimal word pair, an F0 displacement is enough to cause a total identification change, i.e. F0 prevails in a trade-off relation with the energy integral. In the minimal phrase pair, however, we have an extra F0 rise associated with enclitic stress, i.e. F0 is in a trade-off relation with both the word stress F0 rise and the energy integral of the enclitic stress where the latter may be decisive within the enclitic structure. On the other hand, enclitic stress is fixed and its rhythmic function may attribute another status to the energy integral, the invariant acoustic correlate of stress, in comparison with word stress which is rather flexible.

The acoustic parameters seem to contribute more to the perception of stress when combined than when the single parameters' contributions are summed. In the light of this finding, it seems that the idea of the hierarchy of the acoustic cues to stress perception should be reconsidered. A particular acoustic cue whose primary function is to distinguish a phonetic feature may contribute to another phonetic feature as well (Massaro and Oden 1980); in our case, F0, apart from its highlighting function to provide speech elements with the appropriate semantic weighting, may provide the listener with information about the rhythmic structuring in combination with the energy integral, i.e. the perceptual system is predisposed to use any available information. On the other hand, an acoustic cue, for instance intensity, may have minimal perceptual effects in certain environments; however, it may be decisive when combined with the other acoustic parameters in helping listeners make their final linguistic decision.

In the enclitic stress experiment, we could not find out whether F0 manipulations are perceived categorically or continuously (in a broad sense), the way they are with word stress. The explanation may be attributed to the different nature of word and enclitic stress and the different effect F0 has upon them, an all-or-none effect on word stress whereas a moderate one on enclitic stress.

As a partial outcome of this study, we shall disregard sentence stress which has a global manifestation and functions to highlight the most important information. We shall concentrate on enclitic stress and its relation to word stress. Both word stress and enclitic stress have local manifestations within their corresponding domain and both contribute to the rhythmic structure of the language; moreover word stress has a distinctive function at the word level and enclitic stress at the phrase level.

Phonologically, both word and enclitic stress are subjected to the trisyllabic constraint, but its effect is to move word stress to the right at the word level while producing another stress, the enclitic stress, at the phrase level. Acoustically, both word and enclitic stress have the same manifestation but it is enclitic stress which attracts sentence stress regardless of whether word or enclitic stress is in focus within an enclitic structure. Perceptually, F0 may cause a total identification change for word stress but only a moderate one for enclitic stress; furthermore, F0 has categorical tendencies for word stress but not for enclitic stress.

The above observations make it clear that word stress and enclitic stress should not be mixed up in one category dimension with different degrees (Chomsky and Halle 1968) nor seen in a strong ~ weak relation (Liberman and Prince 1977). They are different prosodic categories with partly shared and partly distinct functions. Their common function, rhythmic structuring, may be thought of as a temporal alternation of stressed ~ unstressed syllables with distinction made between stress ~ syllable timed languages (Dauer 1983, Lehiste 1977). Further differentiation of the stressed syllables may be produced on semantic grounds whereas the unstressed syllables may be organized in a strong ~ weak relation (Bruce 1983). However, further research is required on this subject.

5 Summary and Conclusion

5.0 Introduction

The present study is a phonetic investigation of stress and the prosodic structure of Greek and is composed of four parts. In the first part (1), the phonology of Greek prosody is described and the function of the prosodic categories of lexical/word stress, enclitic stress and sentence stress are presented along with rules which describe these categories. The second part (2) is a report of three acoustic experiments; experiment I on the manifestation of word and sentence stress, experiment II on the manifestation of enclitic stress and a comparison of a proclitic structure to an enclitic structure and, experiment III, a report of two different syntactic structures' acoustic manifestations and their relation to prosody. The third part (3) is a report of a physiological investigation of variations of subglottal air pressure (Ps) associated with word and sentence stress. The fourth part (4) is a report of six perceptual experiments, the first four experiments on word stress perception and the last two on enclitic stress perception.

5.1 Phonological Study

In this part of the study, the prosodic categories of Greek are presented and their function and distribution are described. Each grammatical level is associated with a corresponding prosodic category, i.e. the lexical level with lexical stress, the morphological level with word stress, the syntactic level with enclitic stress, and the semantic and textual level with sentence stress.

Both lexical and word stress are subjected to the Trisyllabic Constraint, i.e. stress may not appear to the left of the antepenultimate syllable, as well as to the Monotonic Principle, i.e. a grammatical word - no matter how complex - may have only one stress. Lexical stress is provided by the lexicon for a considerable part of the language's vocabulary and has a distinctive function in words belonging to the same as well as to different grammatical categories. Word stress may coincide with lexical stress, although it usually moves to the right when the word boundary moves to the right by the addition of extra syllables through inflection and derivation or when some morphemes attracting word stress are added to the lexical word.

Enclitic stress appears two syllables to the right of word stress - whether on the lexical element or an enclitic - when the phrase boundary is more than two syllables to the right of word stress; its function is - in combination with lexical/word stress - to organize the rhythmic structuring of the utterance and to realize the syntactic phrasing. Sentence stress appears on the last lexical element bearing focus. However, enclitic stress attracts sentence stress within the domain of the phrase no matter which element is in focus; its function is to highlight the most important information.

The prosodic categories are organized into a hierarchical structure and have a classificatory function; every category has its own distribution rules which apply to the corresponding level. The freedom of the prosodic categories varies according to the level on which their rules are applied. The lower the level the less dependent the prosodic rules are upon the higher levels of representation. In other words, word stress needs information only about the position of lexical stress and the word boundary whereas enclitic stress, apart from word stress information, needs information about phrase boundaries as well, i.e. a higher level is involved. Thus, prosody cannot be independent from the morpholexical, syntactic and semantic structure of the language; it is rather an abstract linguistic structure with distinctive functions, which the different prosodic categories define with concrete contributions from the corresponding levels of representation.

5.2 Acoustic Study

The acoustic part of this study consists of three experiments. In experiment I, the relative contribution of the prosodic parameters of duration, F0 and intensity for word stress in different positions as well as for the manifestation of sentence stress was investigated. In experiment II, the contribution of the prosodic parameters for enclitic stress as well as the prosodic manifestation of a proclitic structure was compared to the one of an enclitic structure. In experiment III, the relation between syntax and prosody was investigated, comparing a paratactic structure carrying two word stresses to an enclitic structure carrying a word and an enclitic stress.

In experiment I, a variable F0 (in an alternating, anathetic or assimilated contour), and constantly longer duration combined with higher relative intensity were found to contribute locally to prefocal word stress, but only duration and intensity to postfocal word stress. An expanded F0 gesture as well as increased duration and intensity contribute to sentence stress locally accompanied by an F0 reorganization of the whole utterance globally. Of the three acoustic parameters examined, duration and intensity go together and are usually present for both word stress and sentence stress whereas F0 may be dissociated from these parameters depending on the context. Thus, F0 appears to be constant in sentence stress and active to a varying degree in non-focal positions, an indication of the basic independent but interrelated character of the acoustic parameters.

In experiment Π, enclitic stress is manifested by a variable F0 rise in prefocal and focal positions and a falling and flattening F0 in postfocal position constantly combined with longer duration and higher peak intensity. The acoustic structure of enclitic stress is basically the same as that of word stress, the difference between the two categories being of perceptual as well as of functional nature. F0 may have an all-or-none effect on word stress whereas it has a partial effect on enclitic stress; on the other hand, word stress has a rhythmic function and is distinctive at the word level, whereas enclitic stress, along with the rhythmic functioning, may be used distinctively at the phrase level. Apart from enclitic stress, the enclitic structure appears with basically the same acoustic manifestation as the proclitic structure although with a weakened word stress.

In experiment III, the influence that syntax may have on prosody has been examined. Two different syntactic structures - a paratactic one with two word stresses and an enclitic one with a word and an enclitic stress - with the same number of syllables, in the same context, may apparently have the same prosodic manifestation out of focus; but when focus is involved, the two syntactic structures may coincide prosodically but may also take completely different prosodic manifestations. This is due to the fact that sentence stress may be applied on any element of the paratactic structure whereas the enclitic structure has a major syntactic constraint, i.e. no matter which element of the enclitic structure with an enclitic stress is in focus, sentence stress is constrained to coincide with enclitic stress.

5.3 Physiological Study

In the physiological part of the study, one experiment investigating the variations of subglottal pressure (Ps) associated with word stress in different positions and with sentence stress was carried out. Moreover, the effect that Ps may have on intensity and F0 was considered. This was possible with an non-invasive method for estimating Ps from records of oral pressure (Po) of a linguistic material in which voiceless stops and vowels alternate.

In an indirect comparison, Ps was found to co-vary with intensity both for sentence stress and word stress in prefocal and postfocal positions as a rule. Ps co-varied with F0 for sentence stress as well as, with some variability, for prefocal word stress whereas it had no effect on F0 in postfocal position. Although there is no EMG data in this experiment, my interpretation of the Ps data is that the larynx is mainly responsible for F0 variations and the subglottal system for intensity. On the other hand, intensity is mainly combined with F0 and partly with duration for sentence stress but only with duration for postfocal word stress, an indication that the acoustic parameters are independent of each other and not produced by the same mechanism.

5.4 Perceptual Study

The perceptual study is a report of six experiments, four on word stress and two on enclitic stress distinctions. Experiment I was to investigate the perceptual salience of duration and intensity for postfocal word stress and experiment Π the perceptual salience of duration, intensity and F0 for prefocal word stress. Experiment ΙΠ was to investigate the perceptual effects of F0 changes on the frequency dimension and experiment IV the perceptual effects of F0 changes on the time dimension for prefocal word stress. Experiment V was to investigate the perceptual salience of duration, intensity and F0 for enclitic stress and experiment VI examined the perceptual effects of F0 changes for the enclitic stress distinction.

In experiment I, duration was found to have a stronger effect than intensity for perception of postfocal word stress, although a combination of both duration and intensity was necessary for the perception of word stress in this position. Listeners could perceive word stress distinctions after focus where the acoustic parameters are weakly represented, even in synthetic speech. The context, F0, and formant structure did not have any considerable perceptual effects in this position.

In experiment II, F0 appeared to have an all-or-none effect on word stress perception in prefocal position. In this position, F0 overrides the word stress acoustic cues of duration and intensity.

In experiment III, it was found that F0 changes along the frequency dimension associated with the prefocal word stress perception had negligible effect on listeners. However, the main finding of this experiment was that the prosodically contrastive words kept their original meaning even with a neutralized F0 contour. Thus, F0 is not an absolute necessity for word stress perception; duration and intensity may equally effectively convey the word stress concept in prefocal position the same way they do in postfocal position.

In experiment IV, the all-or-none effect of F0 on word stress perception was validated. The timing of F0 across the prosodic minimal pair under investigation in prefocal position could cause a total identification change in a rather categorical mode. However, F0 had to be far enough away from the mid-point of the syllable in order to neutralize the influence of the conflicting cues of duration and intensity of the reference word.

In experiment V, listeners could perceive enclitic stress and associate it with the proper syntactic and semantic structure. The enclitic stress F0 rise contributes most to enclitic stress perception, but it has to be combined with duration and intensity to denote the concept of enclitic stress in contrast to word stress in prefocal position, where F0 itself may be the decisive factor.

In experiment VI, F0 changes for the enclitic stress distinction had only a partial effect on listeners. However, the present experiment corroborates the results of experiment V where F0 was not enough to convey enclitic stress by itself, and experiment III where intensity and duration proved to be the decisive perceptual cues when F0 was neutralized.

5.5 Conclusion

To conclude in summary, three prosodic categories have been proposed for the prosodic structure of Greek - word stress, enclitic stress and sentence stress - and their phonetic manifestation as well as their function and their relation to the grammatical levels of the language have been investigated.

The lexicon plays a considerable role for word stress whereas enclitic stress is distributed by rules. Their common function is rhythmic structuring whereas word stress has a distinctive function at the lexical level and enclitic stress may be used for syntactic parsing. Both word stress and enclitic stress are manifested by a combination of duration and intensity, occasionally by F0 in accordance with the requirements of the linguistic context. On the other hand, sentence stress, once the focus domain is given, is distributed by rules. It has a highlighting function and is manifested by a large F0 contribution, typically accompanied by an increase of intensity and duration locally and by a reorganization of the F0 contour of the whole utterance globally.

Although F0 may have an all-or-none effect on listeners, this is thought to be an experimental artefact, which is only indirectly related to word stress perception. F0 is rather supposed to be an intonative parameter the function of which is to attribute to a certain stress group a relative prominence in relation to other stress groups in the utterance; in contrast, duration and intensity, if combined into a single feature, what has been referred to as the energy integral, may be a constant stress correlate the function of which is to organize the rhythmic structure of the utterance. In the light of this argument it seems that the idea about the hierarchy of the acoustic parameters on stress perception should be reconsidered. Primarily, F0 and duration combined with intensity have different functions. On the other hand, a particular parameter (e.g. intensity) may contribute the least by itself; however, it may turn out to be a decisive one when combined with the other acoustic parameters.

Appendices

Appendix 2.1. Five speakers' mean durations (ms), standard deviation (sd) and grand mean ($\bar{\bar{X}}$G) for the minimal pair [ˈmaθima tis~ˈmaθiˈma tis] in the carrier sentence [to ˈneo ___ ˈine ˈðiskolo] in focal position.

(3c~4c)		**ˈm**	**a**	**θ i**	**m**	**a**	t i s	**ˈm**	**a**	**θ** i	**ˈm**	**a**	t i s
1	ms	74	110	100	66	72	204	66	68	108	50	122	214
	sd	5.4	7.0	7.0	5.4	5.4	5.4	4.4	4.4	8.3	0.0	8.3	8.9
2		82	124	94	66	76	186	64	86	96	70	118	196
		8.3	5.4	5.4	5.4	5.4	8.9	5.4	8.9	8.9	7.0	4.4	11.4
3		68	122	94	64	62	196	54	84	90	64	94	192
		4.4	4.4	5.4	5.4	4.4	5.4	5.4	5.4	7.0	5.4	5.4	8.3
4*		50	114	90	52	86	192	44	86	102	54	114	196
		7.0	5.4	7.0	4.4	5.4	4.4	5.4	5.4	8.3	5.4	5.4	5.4
5		64	94	86	44	66	152	54	66	88	54	82	160
		5.4	5.4	8.9	8.9	5.4	8.3	5.4	5.4	4.4	5.4	4.4	7.0
	$\bar{\bar{X}}$G*	67	112	92	58	72	186	59	81	95	59	104	190
		11.9	11.9	5.2	9.9	9.3	20.1	6.4	10.1	9.0	9.1	19.1	22.4

3d~4d)		ˈm	a	θ i	m	a	**t i s**	ˈm	a	θ i	ˈm	a	**t i s**
1	ms	64	92	96	60	66	212	62	90	104	60	112	232
	sd	5.4	4.4	5.4	10.0	5.4	13.0	4.4	0.0	8.9	10.0	4.4	8.3
2		70	102	94	56	72	202	64	92	102	72	110	206
		7.0	8.3	5.4	5.4	8.3	16.4	8.9	8.3	4.4	8.3	7.0	11.4
3		58	86	88	46	60	202	50	84	94	74	106	212
		8.3	5.4	8.3	5.4	7.0	8.3	7.0	5.4	5.4	5.4	5.4	8.3
4*		54	102	80	52	84	192	46	94	102	56	116	214
		5.4	4.4	7.0	4.4	5.4	16.4	5.4	5.4	10.9	5.4	5.4	11.4
5		52	74	74	46	54	156	44	66	94	46	66	154
		4.4	5.4	5.4	5.4	5.4	5.4	5.4	5.4	5.4	5.4	5.4	5.4
	$\bar{\bar{X}}$G	59	91	86	52	67	192	55	83	98	63	98	202
	sd	7.4	11.7	9.3	6.1	11.5	21.7	9.5	11.8	5.2	12.9	21.8	33.2

*Speaker 4 is not included in the calculation of the grand mean ($\bar{\bar{X}}$G) of the enclitic structure.

Appendix 2.2. Five speakers' mean voice fundamental frequency (Hz), standard deviation (sd) and grand mean ($\bar{\bar{X}}$G) of the minimal pair [ˈmaθima tis ~ ˈmaθiˈma tis] in the carrier sentence [to ˈneo ___ ˈine ˈðiskolo], in focal position.

(3c~4c)		**ˈm**	**a**	**θ i**	**m**	**a**	**ˈm**	**a**	**θ i**	**ˈm**	**a**
1	Hz	101	224		101	88	94	102		110	216
	sd	5.4	5.4		7.4	4.4	4.1	10.3		10.0	20.7
2		175	205		122	102	176	206		182	227
		5.0	5.0		2.7	2.7	5.7	4.1		5.7	4.4
3		194	236		143	132	204	185		183	211
		2.7	4.1		4.1	3.5	3.5	4.1		2.7	4.1
4*		168	194		200	186	208	164		143	184
		4.1	5.4		6.8	6.9	5.4	5.0		4.4	2.2
5		126	145		115	103	117	130		127	146
		2.2	4.5		5.7	6.5	2.5	1.0		2.5	3.8
	$\bar{\bar{X}}$G*	148	178		114	101	135	153		143	193
	sd	37.2	45.0		20.6	19.3	36.2	46.5		31.7	36.3

(3d~4d)		ˈm	a	θ i	m	a	ˈm	a	θ i	ˈm	a
1	Hz	101	109		119	126	95	99		106	207
	sd	10.5	14.4		8.9	8.2	3.3	7.5		12.7	12.0
2		154	192		194	191	171	197		189	228
		9.6	2.7		2.2	2.2	8.9	14.8		8.9	17.8
3		136	166		156	142	158	183		148	206
		4.1	4.1		4.1	2.7	8.3	9.7		5.7	5.4
4		158	202		203	194	206	182		180	222
		4.1	2.7		4.4	5.4	5.4	8.3		7.0	4.4
5		121	132		131	128	120	129		122	133
		2.1	2.5		2.1	2.7	0.8	2.6		2.5	2.3
	$\bar{\bar{X}}$G	139	160		160	156	136	152		141	193
	sd	32.0	39.3		37.2	33.7	34.8	45.9		36.2	41.5

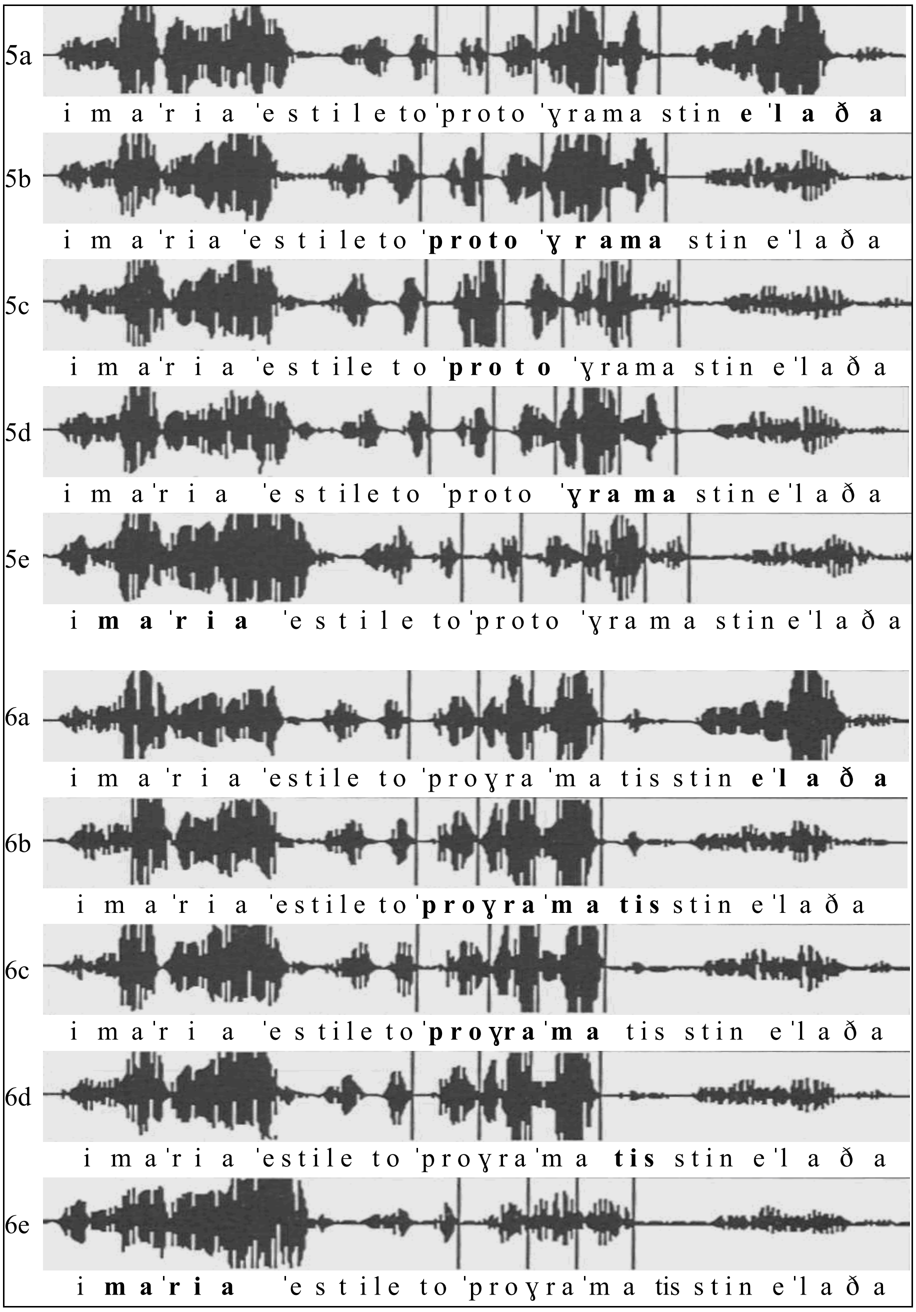
*Speaker 4 is not included in the calculation of the grand mean ($\bar{\bar{X}}$G) of the enclitic structure.

Appendix 2.3. Five speakers' mean intensity (dB), standard deviation (sd) and grand mean ($\overline{\overline{X}}$G) of the minimal pair [ˈmaθima tis ~ ˈmaθiˈma tis] in the carrier sentence [to ˈneo ___ ˈine ˈðiskolo], in focal position.

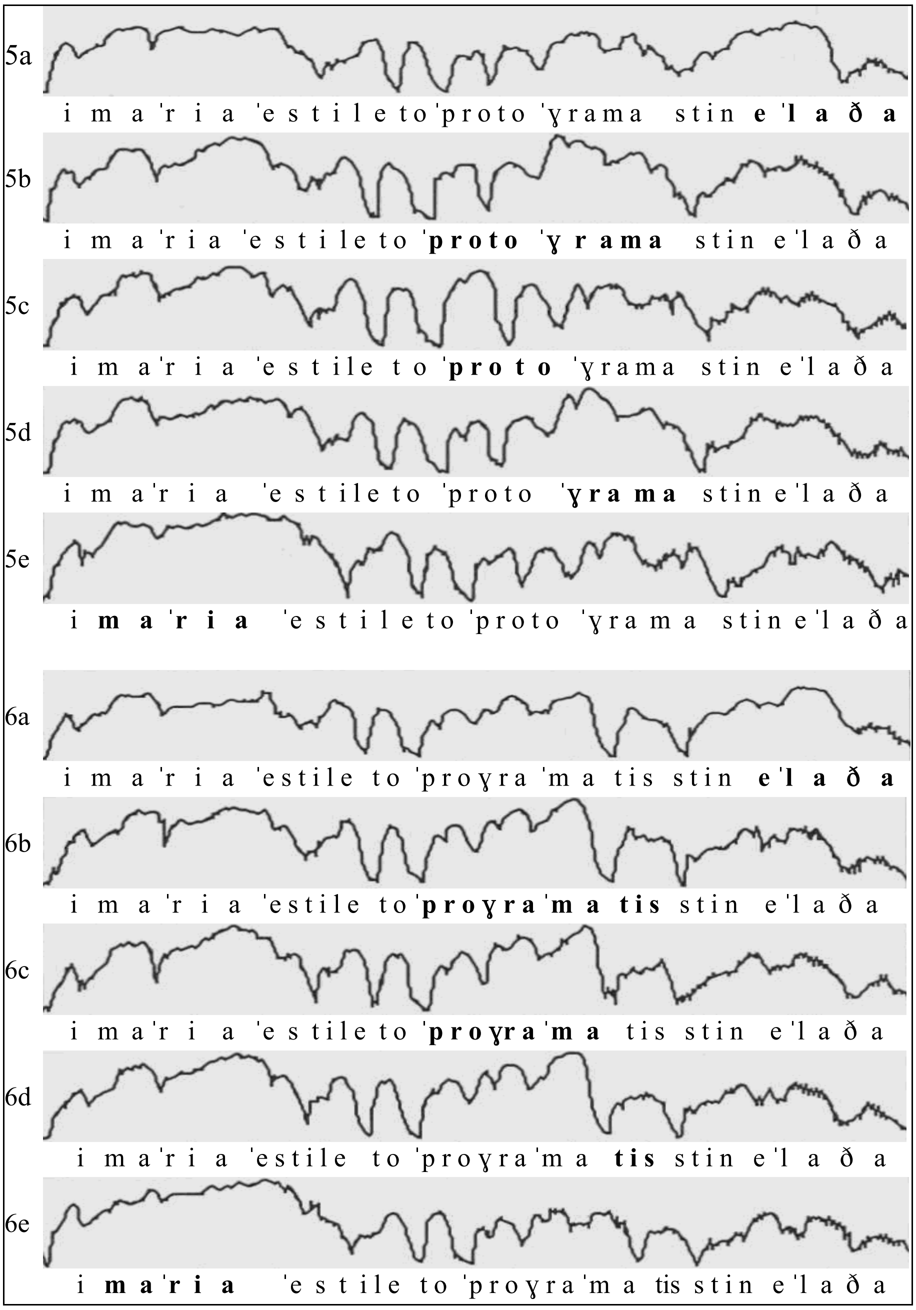
(3c~4c)		**ˈma**	**θi**	**ma**	**ˈma**	**θ i**	**ˈma**
1	dB	-1.0		-18.8	-10.4		-1.0
	sd	0.0		1.0	3.2		0.0
2		-1.0		-16.0	-6.8		-5.4
		0.0		1.4	0.8		1.1
3		-6.8		-19.2	-9.2		-8.4
		1.0		1.0	1.7		1.5
4*		-1.0		-12.0	-4.2		-1.0
		0.0		1.2	1.0		0.0
5		-5.0		-11.4	-9.0		-5.4
		2.1		1.3	1.0		0.5
	* $\overline{\overline{X}}$G	-2.9		-15.4	-8.8		-5.0
		2.7		3.6	1.5		3.0

(3d~4d)		ˈma	θi	ma	ˈma	θi	ˈma
1	dB	-9.6		-14.2	-9.6		-1.0
	sd	3.2		1.3	1.6		0.0
2		-7.6		-13.6	-6.0		-7.4
		0.8		1.3	1.5		1.5
3		-11.2		-16.8	-9.6		-7.4
		1.0		1.0	1.6		0.5
4*		-1.0		-5.2	-2.4		-1.0
		0.0		0.8	0.5		0.0
5		-5.4		-11.0	-5.8		-7.0
		0.5		1.0	0.8		1.0
	$\overline{\overline{X}}$G	-6.9		-12.1	-7.7		-5.7
	sd	3.9		4.4	2.1		3.1

*Speaker 4 is not included in the calculation of the grand mean ($\overline{\overline{X}}$G) of the enclitic structure.



Appendix 2.4. One speaker's durations of the minimal pair [ˈproto ˈɣrama ~ ˈproɣraˈma tis] (5-6) in the carrier sentence [i maˈria ˈestile to __stin eˈlaða] with different focal organizations (a-e).



Appendix 2.5. One speaker's intensity contours of the minimal pair [ˈproto ˈɣrama ~ ˈproɣraˈma tis] (5-6) in the carrier sentence [i maˈria ˈestile to __stin eˈlaða] with different focal organizations (a-e).

Appendix 2.6. Five speakers' mean voice fundamental frequency (Hz), standard deviation (sd) and grand mean ($\bar{\bar{X}}$G) of the minimal pair [ˈproto ˈɣrama ~ ˈproɣraˈma tis] in the carrier sentence [i maˈria ˈestile to ___ stin eˈlaða], in focal position.

(5b~6b)		**ˈp r o**	**t o**	**ˈɣ r a**	**m a**	**ˈp r o**	**ɣ r a**	**ˈm a**	**t i s**
1	Hz	97	105	111	188	95	100	109	187
	sd	2.1	5.0	4.1	8.3	2.5	5.8	10.3	5.7
2		159	189	173	243	164	192	202	249
		4.1	9.6	5.7	4.4	8.2	15.2	14.8	5.4
3		132	160	129	158	131	157	132	160
		5.2	6.9	4.3	2.1	4.1	2.7	5.7	3.5
4		162	198	178	210	158	194	186	209
		2.7	2.7	9.0	9.3	7.5	4.1	4.1	8.9
5		117	136	133	150	113	129	129	150
		4.4	4.1	5.7	5.0	3.7	1.6	1.6	3.5
	$\bar{\bar{X}}$G*	133	157	144	189	132	154	151	191
	sd	27.7	38.2	29.2	38.1	29.2	40.5	40.1	39.8

(5d~6d)		ˈp r o	t o	**ˈɣ r a**	**m a**	ˈp r o	ɣ r a	ˈm a	**t i s**
1	Hz	97	106	116	222	95	100	123	193
	sd	2.5	7.4	9.6	14.8	2.5	4.8	6.7	8.3
2		149	196	167	253	162	194	218	259
		4.1	4.1	10.9	2.7	4.4	8.2	10.9	2.2
3		117	142	111	178	129	158	144	193
		2.7	2.7	4.1	18.2	4.1	11.5	7.4	14.8
4		155	198	170	216	158	196	192	219
		5.0	2.7	7.0	4.1	2.7	4.1	4.4	5.4
5		109	128	124	151	111	131	132	151
		1.0	2.7	4.1	5.4	3.8	2.0	1.7	2.0
	$\bar{\bar{X}}$G	125	154	137	204	131	155	161	203
	sd	25.3	41.3	28.6	39.8	29.1	41.2	41.1	39.6

*Speaker 4 is not included in the calculation of the grand mean ($\bar{\bar{X}}$G) of the enclitic structure.

Bibliography

Abercrombie, D. 1964. Syllable quantity and enclitics in English. In D. Abercrombie, D.B. Fry, P.A.D. McCarthy, N.C. Scott, and G.LM. Trim (eds.). In Honour of Daniel Jones, 216-222. London: Longmans.

Atkinson, J.E. 1978. Correlation analysis of the physiological factors controlling fundamental voice frequency. JASA 63, 211-222.

Baer, T. 1979. Reflex activation of laryngeal muscles by sudden subglottal pressure changes. JASA 65, 1271-1275.

Bannert, R. 1976. Mittelbairische Phonologie ouf akustischer und perzeptorischer Grundlage. Travaux de l'lnstitut de Linguistique de Lund X. Lund: Gleerups.

Bannert, R. 1979. The effect of sentence accent on quantity. Proceedings of the Ninth International Congress of Phonetic Sciences, 253-259. Copenhagen.

Bannert, R. 1982. Temporal and tonal control in German. Working Papers 22, 1-26. Department of Linguistics and Phonetics, Lund University.

Bannert, R. 1985. Towards a model for German prosody. Folia Linguistica XIX/3-4, 321-341.

Bannert, R. 1986. Independence and interdependence of prosodic features. Working Papers 29, 31-60. Department of Linguistics and Phonetics, Lund University.

Bannert, R. and N.G. Thorsen. 1986. Acoustic-Phonetic Investigations of the Intonation of German and Danish: Similarities and Differences. ARIPUC 20, 45-66. Copenhagen.

Beckman, M. E. 1986. Stress and Non-stress Accent. Dordrecht: Foris.

Berinstein, A. E. 1979. A cross linguistic study of perception and production of stress. UCLA, Working Papers in Phonetics, 1-59.

Bloomfield, L. 1933. Language. New York: Holt.

Bolinger, D. L. 1958. A theory of pitch accent in English. Word 14, 109-149.

Bolinger, D. L. 1972. Accent is predictable (if you're a mind reader). Language 48, 633-644.

Botinis, A. 1982. Stress in Modern Greek. Working Papers 22, 27-38. Department of Linguistics and Phonetics, Lund University.

Brame, M.K. 1974. The Cycle in phonology: Stress in Palestinian, Maltese, and Spanish. Linguistic Inquiry 5, 39-60.

Bruce, G. 1977. Swedish Word Accents in Sentence Perspective. Travaux de l'lnstitut de Linguistique de Lund XII. Lund: Gleerups.

Bruce, G. 1981. Tonal and temporal interplay. Working Papers 21, 49-60. Department of Linguistics and Phonetics, Lund University.

Bruce, G. 1982. Developing the Swedish intonation model. Working Papers 22, 51-116. Department of Linguistics and Phonetics, Lund University.

Bruce, G. 1983. Accentuation and timing in Swedish. Folia Linguistica XVII/1-2, 221-238.

Bruce, G. 1984. Rhythmic alternation in Swedish. In C.C. Elert, I. Johansson and E. Strangert (eds.). Nordic Prosody IΠ, 31-41. Umeå Studies in the Humanities 59.

Bruce, G. 1985. Structure and functions of prosody. In Guerin and Carre (eds.). Proceeding of the French Swedish Seminar on Speech, 549-559. Grenoble.

Bruce, G. and E. Gårding. 1978. A prosodic typology for Swedish dialects. In E. Gårding, G. Bruce and R. Bannert (eds.). Nordic Prosody, 219-228. Travaux de l'lnstitut de Linguistique de Lund XIΠ. Lund: Gleerups.

Chomsky, N. 1970. Deep structure, surface structure, and semantic interpretation. In D. Steinberg and L. Jakobovits (eds.). Semantics: An Interdisciplinary Reader in Philosophy, Linguistics and Psychology, 183-216. Cambridge: Cambridge University Press.

Chomsky, N. and M. Halle. 1968. The Sound Pattern of English. New York: Harper & Row.

Collier, R. 1975. Physiological correlates of intonation patterns. JASA 58, 249-255.

Cooper, W.E. and J.M. Sorensen. 1981. Fundamental Frequency in Sentence production. New York: Springer-Verlag.

Cooper, W.E., S.J. Eady and P.R. Mueller. 1985. Acoustical aspects of contrastive stress in question-answer contexts. JASA 77, 2142-2156.

Couper-Kuhlen, E. 1986. An Introduction to English Prosody. London: Edward Arnold.

Crystal, D. 1969. Prosodic Systems and Intonation in English. Cambridge: Cambridge University Press.

Danes, F. 1960. Sentence intonation from a functional point of view. Word 16, 34-54.

Danes, F. 1967. Order of elements and sentence intonation. In To Honour Roman Jakobson, 499-512. The Hague: Mouton.

Danes, F. 1974. Functional sentence perspective and the organisation of the text. In F. Danes (ed.). Papers on Functional Sentence Perspective, 106-128. Janu Linguarum, Series Minor 147.

Dauer, R.M. 1983. Stress-timing and syllable-timing reanalysed. Journal of Phonetics 11, 51-62.

Di Cristo, A. 1978. De la Microprosodie a l'lntonosyntaxe. Thèse de Doctorat d'Etat. Universitè de Provence.

Di Cristo, A. 1981a. L' intonation est congruente a la syntaxe: Une confirmation. In M. Rossi, A. Di Cristo, D. Hirst, Ph. Martin and Y. Nishinuma. L'Intonation. De l'Acoustique a la Sémantique, 272-289. Paris: Klincksieck.

Di Cristo, A. 1981b. Aspects phonétiques et phonologiques des elements prosodiques. Modeles Linguistiques III, 24-83. Lille: Presses Universitaires.

Eady, S.J., W.E. Cooper, G.V. Klouda, P.R. Mueller and D.W. Lotts. 1986. Acoustical characteristics of sentential focus: Narrow vs. broad and single vs. dual focus environments. Language and Speech 29, 233-251.

Farnetani, E. and S. Kori. 1986. Effects of syllable and word structure on segmental durations in spoken Italian. Speech Communication 5, 17-34.

Firbas, J. 1974. Some aspects of the Czechoslovak approach to problems of functional sentence perspective. In F. Danes (ed.). Papers on Functional Sentence Perspective, 11-37. Janu Linguarum, Series Minor 147.

Fisher-Jørgensen, E. 1984. The acoustic manifestation of stress in Danish with particular reference to the reduction of stress in compounds. ARIPUC 18,45-161. Copenhagen.

Fonagy, I. 1966. Electrophysical and acoustic correlates of stress and stress perception. Journal of Speech and Hearing Research 9, 231-244.

Fourakis, M. 1986. An acoustic study of the effects of tempo and stress on segmental intervals in Modern Greek. Phonetica 43, 172-188.

Fry, D.B. 1955. Duration and intensity as physical correlates of linguistic stress. JASA 27, 765-768.

Fry, D.B. 1958. Experiments in the perception of stress. Language and Speech 1, 126-152.

Fry, D.B. 1965. The dependence of stress judgements on vowel formant structure. In Proceedings of the Fifth International Congress of Phonetic Sciences, 306-311. Minister.

Garde, P. 1968. L' Accent. Paris: Press Universitaires de France.

Gårding, E. 1967. Prosodiska drag i spontant och uppläst tal. In G. Holm (ed.). Svensk talspråk, 40-85. Stockholm: Almqvist & Wiksell.

Gårding, E. 1973. The Scandinavian Word Accents. Working Papers 8. Department of Linguistics and Phonetics, Lund University.

Gårding, E. 1977a. The Scandinavian Word Accents. Travaux de l'lnstitut de Linguistique de Lund XI. Lund: Gleerups.

Gårding, E. 1977b. The importance of turning-points for the pitch patterns of Swedish accents. In L. M. Hyman (ed.). Studies in Stress and Accent. Los Angeles: Southern California Occasional Papers in Linguistics 4, 27-35.

Gårding, E. 1984. Comparing intonation. Working Papers 27, 75-99. Department of Linguistics and Phonetics, Lund University.

Gårding, E. 1987. How many intonation models are there in Lund? Working Papers 31, 1-9. Department of Linguistics and Phonetics, Lund University.

Gårding, E., O. Fujimura and H. Hirose. 1970. Laryngeal control of Swedish word tone: A preliminary report on an EMG study. Annual Bulletin, Research Institute of Logopedics and Phoniatrics 4, 45-54. University of Tokyo.

Gårding, E. and G. Bruce. 1981. A presentation of the Lund model for Swedish intonation. In T. Fretheim (ed.). Nordic Prosody II, 33-39. Trondheim: Tapir.

Gårding, E., A. Botinis and P. Touati. 1982. A comparative study of Swedish, Greek and French intonation. Working Papers 22, 137-153. Department of Linguistics and Phonetics, Lund University.

Gussenhoven, C. 1983. Focus, mode and the nucleus. Journal of Linguistics 19, 377-417.

Hadding-Koch, K. 1961. Acoustico-phonetic Studies in the Intonation of Southern Swedish. Travaux de l'lnstitut de Linguistique de Lund III. Lund: Gleerups.

Hadding-Koch, K. and M. Studdert-Kennedy 1964. An experimental study of some intonation contours. Phonetica 11, 175-184.

Halliday, M. A. K. 1967. Intonation and Grammar In British English. The Hague: Mouton.

t'Hart, J. and A. Cohen. 1973. Intonation by rule: a perceptual quest. Journal of Phonetics 1, 309-327.

Hayes, B. 1984. The phonology of rhythm in English. Linguistic Inquiry 15, 33-74.

Hirst, D. 1981. Le modele de production. In M. Rossi, A. Di Cristo, D. Hirst, Ph.

Martin and Y. Nishinuma. L'Intonation. De I 'Acoustique a la Sémantique, 30-39. Paris: Klincksieck.

Hockett, C.F. 1942. A system of descriptive phonology. Language 18, 3-21.

Holmberg, E. 1980. Laryngeal airway resistance as a function of phonation type. JASA 68, S101 (A).

Home, M. 1985.English sentence stress, grammatical functions and contextual coreference. Studia Linguistica 39,51-66.

Huss, V. 1978. English word stress in the post-nuclear position. Phonetica 35, 86-105.

Hyman, L. M. 1977. On the nature of linguistic stress. In L.M. Hyman (ed.). Studies in Stress and Accent, 37-82. Los Angeles: Southern California Occasional Papers in Linguistics 4.

Isshiki, N. 1964. Regulatory mechanism of voice intensity variation. Journal of Speech and Hearing Research 7,17-29.

Jackendoff, R. S. 1972. Semantic Interpretation in Generative Grammar. Cambridge, Mass.: MIT Press.

Jones, D. 1950. The Phoneme: Its Nature and Use. Cambridge: W. Heffner & Sons.

Kahane, H. and R. Ward. 1945. Spoken Greek. New York: Henry Holt and Company.

Kean, M. L. 1974. The strict cycle in phonology. Linguistic Inquiry 5, 179-203.

Kiparsky, P. 1979. Metrical structure assignment is cyclic. Linguistic Inquiry 10, 421-441.

Klatt, D.H. 1976. Linguistic use of segmental duration in English: acoustic and perceptual evidence. JASA 59, 1208-1221.

Klavans, J.L. 1985. The independence of syntax and phonology in cliticization. Language 61, 95-120.

Koutsoudas, A. 1962. Verb Morphology of Modern Greek. Bloomington: Indiana University Press.

Ladd, D. R. 1980. The Structure of Intonational Meaning. Bloomington: Indiana University Press.

Ladefoged, P. 1967. Three Areas of Experimental Phonetics. London: Oxford University Press.

Ladefoged, P. and N. P. McKinney. 1963. Loudness, sound pressure, and subglottal pressure in speech. JASA 35, 454-461.

Lea, W.A. 1977. Acoustic correlates of stress and juncture. In L. M. Hyman

(ed.). Studies in Stress and Accent, 88-119. Los Angeles: Southern California Occasional Papers in Linguistics 4.

Lea, W.A. 1980. Prosodic aids to speech recognition. In W. A. Lea (ed.). Trends in Speech Recognition, 166-205. N. L: Prentice-Hall.

Lehiste, I. 1970. Suprasegmentals. Cambridge, Mass.: M.I.T. Press.

Lehiste, I. 1973. Phonetic disambiguation of syntactic ambiguity. Glossa 7, 107-122.

Lehiste, I. 1977. Isochrony reconsidered. Journal of Phonetics 5, 253-263.

Liberman, M. and A. Prince. 1977. On Stress and Linguistic Rhythm. Linguistic Inquiry 8, 249-336.

Liberman, M. and J. Pierrehumbert. 1984. Intonational invariance under changes in pitch range and length. In M. Aranoff & R. T. Oehrle (eds.). Language Sound Structure: Studies in Phonology Presented to Morris Halle. MIT Press.

Lieberman, Ph. 1960. Some acoustic correlates of word stress in American English. JASA 32, 451-454.

Lieberman, Ph. 1967. Intonation, Perception, and Language. Cambridge, Mass.: MΓT Press.

Lindblom, B. and K. Rapp. 1973. Some Temporal Regularities of Spoken Swedish. PILUS 21. Stockholm.

Lindsay, D. and W.A. Ainsworth. 1985. Two modes of nuclear intonation. Journal of Phonetics 13, 163-173.

Lofqvist, A., B. Carlborg and P. Kitzing 1982. Initial validation of an indirect measure of subglottal pressure during vowels. JASA 72, 633-635.

Lyberg, B. 1981. Some observations on the vowel duration and the fundamental frequency contour in Swedish utterances. Journal of Phonetics 9, 261-272.

Malikouti-Drachman, A. 1976. On the stress cycle in Greek. Salzburger Beitrage zur Linguistik 1, 105-123.

Malmberg, B. 1955. Observations on the Swedish word accent. Haskins Laboratories mimeographed report.

Martinet. A. 1965. La linguistique synchronique. Paris: Presses Universitaires de France.

Martinet, A. 1960. Elements de Linguistique Generate. Paris: Armand Colin.

Massaro, D.W. and G.C. Oden. 1980. Evaluation and integration of acoustic features in speech perception. JASA 67, 996-1012.

Morton, J. and W. Jassem. 1965. Acoustic correlates of stress. Language and Speech 8, 159-181.

Ohala, J.J. 1977. The physiology of stress. In L. Hyman (ed.). Studies in Stress and Accent, 145-168. Southern California Occasional Papers in Linguistics 4

Ohala, J.J. 1978. Production of tone. In V. Fromkin (ed.). Tone: A Linguistic

Survey. 5-41. New York: Academic Press.

Ohman, S.E.G., S. Zetterlund, L. Nordstrand and O. Engstrand. 1979. Predicting segment durations in terms of a gesture theory of speech production. Proceedings of the Ninth International Congress of Phonetic Sciences, 305-311. Copenhagen.

O'Shaughnessy, D. 1979. Linguistic features in fundamental frequency patterns. Journal of Phonetics 7, 119-145.

Paccia-Cooper, J. and W.E. Cooper. 1981. The processing of phrase structures in speech production. In R.A. Cole (ed.). Perception and Production of Fluent Speech, 399-440. Hillsdale, N. J.: Lawrence Erlbaum.

Pierrehumbert, J. 1979. The perception of fundamental frequency declination. JASA 66, 363-369.

Pierrehumbert, J. 1980. The Phonology and Phonetics of English Intonation. Ph. D. Dissertation. MIT.

Pike, K.L. 1945. The Intonation of American English. Ann Arbour: University of Michigan.

Prince, A. 1983. Relating to the grid. Linguistic Inquiry 14, 19-100.

Rischel, J. 1983. On unit accentuation in Danish - and the distinction between deep and surface phonology. Folia Linguistica XVII/1-2, 51-99.

Rossi, M. 1981. Intonation, enonciation, syntaxe. In M. Rossi, A. Di Cristo, D. Hirst, Ph. Martin and Y. Nishinuma. L'Intonation. De l'Acoustique a la Sémantique, 184-233. Paris: Klincksieck.

Rossi, M. 1985. L'intonation et l'organisation de l' énoncé. Phonetica 42, 135-153.

Rossi, M., A. Di Cristo, D. Hirst, Ph. Martin and Y. Nishinuma. 1981. L'Intonation. De l'Acoustique a la Sémantique. Paris: Klincksieck.

Schane, S.A. 1979. Rhythm, accent, and stress in English words. Linguistic Inquiry 10, 483-502.

Selkirk, E. 1980. The role of prosodic categories in English word stress. Linguistic Inquiry 11, 563-605.

Smitheran, J. R. and T. J. Hixon. 1981. A clinical method for estimating laryngeal airway resistance during vowel production. Journal of Speech and Hearing Disorders 46,138-146.

Sotiropoulos, D. 1972. Noun Morphology of Modern Greek. The Hague: Mouton.

Stathopoulos, E.T. and G. Weisner. 1985a. An aerodynamic study of stress in children and adults. Journal of Phonetics 13, 343-355.

Stathopoulos, E.T. and G. Weisner. 1985b. Effects of monitoring vocal intensity on oral air flow in children and adults. Journal of Speech and Hearing Research 28, 589-593.

Steriade, D. 1988. Greek Accent: A case for preserving structure. Linguistic Inquiry 19, 271-314.

Strangert, E. 1985. Swedish Speech Rhythm in a Cross-Language Perspective. Umeå Studies in the Humanities 69. Umeå.

Studdert-Kennedy, M., A.M. Liberman, K.S. Harris, and F.S. Cooper. 1970. The motor theory of speech perception: A reply to Lane's critical review. Psychological Review 77, 234-249.

Studdert-Kennedy, M. and K. Hadding-Koch. 1973. Auditory and linguistc processes in the perception of intonation contours. Language and Speech 16, 293-313.

Tanaka, S. and W.J. Gould. 1983. Relationships between vocal intensity and noninvasively obtained aerodynamic parameters in normal subjects. JASA 73, 1316-1321.

Tevajarvi, K. 1982. Intonation in Finland Swedish: Word and sentence stress in

the Helsinki dialect. Working Papers 22, 175-180. Department of Linguistics and Phonetics, Lund University.

Thorsen, N. 1978. An acoustic investigation of Danish intonation. Journal of Phonetics 6, 151-175.

Thorsen, N. 1980. Neutral stress, emphatic stress, and sentence intonation in Advanced Standard Copenhagen Danish. ARIPUC 14, 121-205. Copenhagen.

Thorsen, N. 1982. On the variability in F0 patterning and the function of F0 timing in language where pitch cues stress. Phonetica 39, 302-316.

Thorsen, N. 1983. Standard Danish sentence intonation - phonetic data and their representation. Folia Linguistica 17, 187-220.

Touati, P. 1987. Structure Prosodiques du Suedois et du Frangais. Travaux de L'lnstitut de Linguistique de Lund 21. Lund University Press.

Trager, G.L. and H.L. Smith. 1951. Outline of English Structure. Studies in Linguistics No 3. Norman, Okla.: Battenburg Press.

Trubetskoy, N.S. 1969. Principles of Phonology. Translated by C. Baltaxe. University of California Press.

Tsitsopoulos, S. 1973. Stress in Modern Greek. Ph. D. Dissertation. Univiversity of Michigan.

Vanderlice, R. 1967. Larynx vs lungs: Cricothyrometer data refuting some recent claims concerning intonation and archetypality. UCLA, Working Papers in Phonetics 7, 69-79.

Vanderslice, R. and P. Ladefoged. 1972. Binary suprasegmental features and transformational word-accentuation rules. Language 48, 819-838.

van Katwijk, A. 1974. Accentuation in Dutch: An Experimental Study. Ph. D. Dissertation, Utrecht University.

Warburton, I. 1970a. On the verb in Modern Greek. Indiana University, Bloomington. The Hague: Mouton (Language Science Monographs Volume 4).

Warburton, I. 1970b. Rules of accentuation in Classical and Modern Greek. Glotta 48,107-120.

Williams, B. 1986. An acoustic study of some features of Welsh prosody. In C. Johns-Lewis (ed.). Intonation in Discourse, 35-51. London: Croom Helm.

Zwicky, A. 1977. On clitics. Bloomington: Indiana University Linguistics Club.

Zwicky, A. 1985. Clitics and particles. Language 61, 283-305.

Name index

Language index

TRAVAUX DE L’INSTITUT DE LINGUISTQUE DE LUND

FONDÉS PAR BERTL MALMBERG
PUBLIÉS PAR GÖSTA BRUCE ET BENGT SIGURD

1 Carl-Gustav Söderberg
A Typological Study on the Phonetic Structure of English Words with an Instrumental-Phonetic Excursus on English Stress. 1959.
2 Peter S. Green
Consonant-Vowel Transitions. A Spectrographic Study. 1959.
3 Kerstin Hadding-Koch
Acoustic-Phonetic Studies in the Intonation of Southern Swedish. 1961.
4 Börje Sederbäck
La realization däune opposition de tonèmes dans des dissyllabes chuchotés. Étude de phonétique expérimentale. 1966.
5 Velta Rūķe-Draviņa
Mehrsprachigkeit im Vorschulalter. 1967.
6 Eva Gårding
Internal Juncture in Swedish. 1967.
7 Folke Strenger
Les voyelles nasales françaises. 1969.
8 Edward Carney
Hiss Transitions and their Perception. 1970.
9 Faith Ann Johansson
Immigrant Swedish Phonology. 1973.
10 Robert Bannert
Mittelbairische Phonologie auf akustischer und perzeptorischer Grundlage. 1976.
11 Eva Gårding
Internal Juncture in Swedish. 1967.
12 Gösta Bruce
Swedish Word Accents in Sentence Perspective. 1977
13 Gösta Bruce, Eva Gårding, Robert Bannert (eds.)
Nordic Prosody. 1978.
14 Ewa Söderpalm
Speech errors in Normal and Pathological Speech. 1979.
15 Kestin Nauclér
Perspectives on Misspellings. 1980.
16 Per Lindbland
Svenskans sje- och the-ljud (some Swedish sibilants). 1980.
17 Eva Magnusson
The Phonology of Language Disordered Children. 1983.
18 Jan-Olof Svantesson
Kammu Phonology and Morphology. 1983.
19 Ulrika Nettelbladt
Developmental Studies of Disphonology in Children. 1983.
20 Gisela Håkansson
Teacher Talk. How teachers modify their speech when addressing learners of Swedish as a foreign language. 1987.
21 Paul Touati
Structures prosodiques du suédois et du français. Profils temporels et configurations tonales. 1987.
22 Antonis Botinis
Stress and Prosodic Structure in Greek. A Phonological, Acoustic, Physiological and Perceptual Study. 1989.

Antonis Botinis

Stress and Prosodic Structure in Greek
A Phonological, Acoustic, Physiological and perceptual Study

This is phonetic study of stress and prosodic structure in Greek. Prosody is how we use temporal, dynamic and tonal variations in speech to convey linguistic (as well as paralinguistic) information. Word prosody is used to distinguish words with different meanings (morphological function); phrase and sentence prosody are used to group lexical elements together (syntactic function) and to divide the utterance into more or less prominent parts in accordance with the speaker's/listener's idea as to what has greater or lesser importance (semantic function).

The study falls naturally into four parts. After a short introduction, the first part of the study outlines the prosodic phonology of Greek, and rules describing the distribution of the prosodic categories of word, enclitic and sentence stress are introduced. In the second part, the acoustic correlates of duration, intensity and voice fundamental frequency for word, enclitic and sentence stress are investigated as well as the relation between prosody and syntax/semantics. In the third part, subglottal pressure, as a physiological correlate of word and sentence press, and its relation to the acoustic correlates of word and sentence stress is investigated. Finally, in the fourth part, the perceptual correlates of word and enclitic stress are examined and the perceptual relevance of voice fundamental frequency for stress distinctions is discussed at length. At the end, the study is summarized and conclusions are presented.

TRAVAUX DE L'INSTITUT DE LINGUISTIQUE DE LUND

Lund University Press
Box 141. S-221 00 Lund, Sweden
Art nr 20137
ISSN 0347-2558
ISBN 91-7966-080-0

Chartwell-Bratt
Old Orchard, Bickley Road,
Bromley, Kent BR1 2NE
England
ISBN 0-86238-223-8