THE PHONETICS OF MODERN GREEK RHYTHM AND ITS PHONOLOGICAL IMPLICATIONS

Amalia Arvaniti
Clare Hall, Cambridge

A dissertation submitted to the University of Cambridge
for the degree of Doctor of Philosophy.

Department of Linguistics
11th April 1991

SUMMARY

Speech rhythm has been the subject both of phonetic and of phonological studies. However, neither exclusively phonetic nor exclusively phonological studies have achieved an adequate description of speech rhythm. The phonetic distinction into stress- and syllable-timed languages has not been supported by the results of acoustic studies. Despite the lack of empirical evidence for the two rhythmic categories, metrical phonology has adopted the distinction and incorporated it into metrical representation. Moreover, the principles of metrical phonology, which was originally intended as a description of English rhythm, have been adopted uncritically for the rhythmic description of widely different languages, such as French, Polish and Greek.

In contrast, the present thesis adopts an interdisciplinary approach to the study of rhythm. The Greek metrical structure is investigated by means of perceptual experiments and acoustic analyses of natural speech. The main topics under study are the acoustic correlates of primary and rhythmic stress and the perceptual and acoustic nature of the two stresses in a host-and-clitic group with antepenultimate stress. The experimental results form the basis on which a phonological representation of Greek rhythm is elaborated.

The empirical data show that the acoustic correlates of primary stress in Greek are amplitude, duration and fundamental frequency. Concerning the stresses in a host-and-clitic group the following picture emerges: the enclitic (added) stress is indirectly shown to be more prominent than the original lexical stress of the host (secondary stress); the secondary stress, which has often been equated to rhythmic stress, is perceptually and acoustically distinct from rhythmic stress and identical to a lexical stress which is not the designated terminal element (i.e. the head) of the phonological phrase it belongs to. No convincing acoustic evidence for rhythmic stress was found.

The above results suggest that the rhythmic structure of Greek is based chiefly on stress and is flatter and more flexible than has been previously assumed, as lapses and clashes are tolerated to a considerable extent. The rhythmic features of Greek suggest, first, that accent (in Greek manifested as stress) is the main contributor to rhythm, irrespective of the rhythmic category a language is thought to belong to; second, that binary rhythmic patterns are not universal as phonologists have often assumed. Thus, it is proposed that the stress-/syllable-timing classification be replaced by an abstract phonological representation, which reflects the contribution of accent to speech rhythm. This representation should be adapted to the particular rhythmic features of each language. It is proposed that Greek rhythm is best represented by n-ry branching trees which have only 5 levels: the syllable, the phonological word, the phonological phrase, the intonational phrase and the phonological utterance.

DECLARATION

This thesis includes nothing which is the result of work done in collaboration with others, and is original save where explicit reference is made in the text to other works. No part of this thesis has been submitted in any form for any qualification at any other University.

The length of this thesis does not exceed 65,000 words including footnotes and references.

ACKNOWLEDGMENTS

I have here the opportunity to thank all those who contributed to the preparation of this thesis. First of all, I would like to thank my parents; they financed me ungrudgingly throughout my postgraduate studies and with my brother provided me with a safe resting place back home.

Academically this thesis owes much to discussions with several people. As they have not always been able to change my mind, the blame for all errors and omissions remains with me. I would like to thank first of all my supervisor, Sarah Hawkins, who has offered me tireless assistance in every way, perhaps most of all by constantly questioning my work; her comments forced me to clarify my ideas. I am also indebted to other members of the Linguistics Department at Cambridge. Terry Moore was the first person to give me confidence and I warmly thank him for that. Francis Nolan has been very helpful, offering me enlightening suggestions and moral support in critical moments. My friendly arguments with April McMahon have been very useful in crystallising my ideas on theoretical questions.

My year in Edinburgh was a happy period, from the point of view of this thesis. I would like to thank most of all Bob Ladd whose enthusiasm, support and extensive comments on my work have been invaluable. I am also grateful to Jan McAllister, Ellen Bard, Irene MacLeod and Ced MacMartin for their help with statistical and technical aspects of the work. Outside Cambridge and Edinburgh, Marina Nespor, Mary Beckman and Brian Joseph have also contributed with their comments on and support for my work in the preparation of this thesis.

Even my parents' generous financial help would have been insufficient without the assistance of Clare Hall and the British Council. I am particularly indebted to Clare Hall for offering me the opportunity to start my postgraduate studies in Cambridge and for helping me with fees and conference expenses for a number of years. I am also grateful to the British Council for offering me a generous Fellowship in 1988-89.

I would also like to thank my friends in Cambridge and Edinburgh who provided me with the necessary recreation when thesis preparation and writing became hard to bear. Last but not least, I would like to thank Kimon Roussopoulos who patiently proof-read several drafts of this thesis, gave his assistance with technical aspects of it and offered me moral support throughout the years.

CONTENTS

CHAPTER 1: INTRODUCTION

CHAPTER 2: THE ACOUSTICS OF GREEK STRESS

CHAPTER 3: THE PERCEPTION AND ACOUSTICS OF ENCLITIC STRESS

CHAPTER 6: THE PHONOLOGY OF GREEK RHYTHM

LIST OF TABLES

CHAPTER 2: THE ACOUSTICS OF GREEK STRESS

LIST OF FIGURES

CHAPTER 2: THE ACOUSTICS OF GREEK STRESS

CHAPTER 3: THE PERCEPTION AND ACOUSTICS OF ENCLITIC STRESS

CHAPTER 4: RHYTHMIC AND SECONDARY STRESS

CHAPTER 5: THE ACOUSTICS OF RHYTHMIC STRESS

LIST OF ABBREVIATIONS

acc. = accusative
adj. = adjective
AI = amplitude integral
ANOVA = analysis of variance
C = clitic group
c. ω = compound phonological word
colloq. = colloquial
CV = consonant-vowel
DTE = designated terminal element
F0 = fundamental frequency
fem. = feminine
gen. = genitive
I = intonational phrase
imp. = imperative
IPA = International Phonetic Association
lit. = literally
LP = Liberman & Prince (1977)
mas. = masculine
M-DD = Malikouti-Drachman & Drachman (1980)
met. = metaphorically
MID = Mirror Image Deletion
n. = noun
NA = Nasal Assimilation
ND = Nasal Deletion
neu. = neutral
NV = Nespor & Vogel (1986) and/or (1989)
nom. = nominative
PA = peak amplitude
pers. comm. = personal communication
pl. = plural
RMS = root mean square (average amplitude)
RPR = Right Prominence Rule
RS words = rhythmic stress test-words (Chapter 4)
[s] = strong constituent
sing. = singular
SLH = Strict Layer Hypothesis
SPE = The Sound Pattern of English (Chomsky & Halle 1968)
SS words = secondary stress test-words (Chapter 4)
SV = Stop Voicing
SWFC = Stress Well-Formedness Condition
U = phonological utterance
VOT = voice onset time
[w] = weak constituent
z = minimal phrase (Kondoravdi 1990)

σ = syllable
Σ = foot
ϕ = phonological phrase
ω = phonological word

' = primary stress on the following syllable
. = secondary stress (Chapters 3 & 4) or rhythmic stress (Chapter 5) on the following syllable

To my parents

ΙΘΑΚΗ

Σα βγεις στον πηγαιμό για την Ιθάκη,
να εύχεσαι νάναι μακρύς ο δρόμος,
γεμάτος περιπέτειες, γεμάτος γνώσεις.
Τους Λαιστρυγόνας και τους Κύκλωπας,
τον θυμωμένο Ποσειδώνα μη φοβάσαι,
τέτοια στον δρόμο σου ποτέ σου δεν θα βρεις,
αν μεν΄ η σκέψις σου υψηλή, αν εκλεκτή
συγκίνησις το πνεύμα και το σώμα σου αγγίζει.
Τους Λαιστρυγόνας και τους Κύκλωπας,
τον άγριο Ποσειδώνα δεν θα συναντήσεις,
αν δεν τους κουβανείς μες στην ψυχή σου,
αν η ψυχή σου δεν τους στήνει εμπρός σου.

Να εύχεσαι νάναι μακρύς ο δρόμος.
Πολλά τα καλοκαιρινά πρωϊά να είναι
που με τι ευχαρίστηση, με τι χαρά
θα μπαίνεις σε λιμένας πρωτοειδωμένους·
να σταματήσεις σ΄ εμπορεία Φοινικικά,
και τες καλές πραγμάτειες ν΄ αποκτήσεις,
σεντέφια και κοράλλια, κεχριμπάρια κ΄ εβένους,
και ηδονικά μυρωδικά κάθε λογής,
όσο μπορείς πιο άφθονα ηδονικά μυρωδικά·
σε πόλεις Αιγυπτιακές πολλές να πας,
να μάθεις και να μάθεις απ΄ τους σπουδασμένους.

Πάντα στον νου σου νάχεις την Ιθάκη.
Το φθάσιμον εκεί ειν΄ ο προορισμός σου.
Αλλά μη βιάζεις το ταξείδι διόλου.
Καλλίτερα χρόνια πολλά να διαρκέσει·
και γέρος πια ν΄ αράξεις στο νησί,
πλούσιος με όσα κέρδισες στον δρόμο,
μη προσδοκώντας πλούτη να σε δώσει η Ιθάκη.

Η Ιθάκη σ έδωσε τ΄ ωραίο ταξείδι.
Χωρίς αυτήν δεν θάβγαινες στον δρόμο.
Άλλα δεν έχει να σε δώσει πια.

Κι αν πτωχική την βρεις, η Ιθάκη δεν σε γέλασε.
Έτσι σοφός που έγινες, με τόση πείρα,
ήδη θα το κατάλαβες η Ιθάκες τι σημαίνουν.

Κ.Π. Καβάφης 1911

ITHAKA

When setting out upon your way to Ithaca,
wish always that your course be long,
full of adventure, full of lore.
Of the Laestrygones and of the Cyclops,
of the irate Poseidon don't be afraid;
you'll never find such things on your way
if lofty is your thinking, if fine sentiment
in spirit and in body touches you.
Neither Laestrygones nor Cyclops,
nor wild Poseidon will you ever meet,
unless you bear them in your soul,
unless your soul sets them up in front of you.

Wish always that your course be long;
that there be many summer morns
when, with what pleasure, what joy,
you enter ports now for the first time seen;
that you may stop at some Phoenician marts,
to purchase there the best of wares,
mother of pearl and coral, amber, ebony,
sensual perfumes of all sorts-
as many sensual perfumes as you can;
that you may visit many Egyptian cities,
to learn and learn from their scholars.

Keep Ithaca always in your mind.
Arriving there is your goal.
But do not rush your journey at all.
Better that it should last for many years,
and that, old, you moor at Ithaca at last,
wealthy with all you've gained on the way,
not expecting Ithaca to give you riches.

Ithaca gave you the marvellous journey.
Without her you wouldn't have set out.
She has nothing left to give you now.

And if you find her poor, Ithaca won't have fooled you.
As wise as you have become, of such experience,
you must already understand what Ithacas mean.

Translation based on those of Kolaitis and Keely & Sherrard[1]

---

1 *The Greek Poems of C.P. Cavafy: As translated by Memas Kolaitis*. New York: A.D. Caratzas, Publisher. 1989. C.P. Cavafy. *Collected Poems*. Translated by Edmund Keely & Philip Sherrard. London: Chatto & Windus. 1975.

1. INTRODUCTION

1.1 PRELIMINARY REMARKS

The aim of the present thesis is to report on a study of the phonology and phonetics of rhythm in Modern Greek (henceforth Greek). A two-fold study of rhythm that takes into account both phonology and phonetics has become increasingly necessary as neither exclusively phonetic nor exclusively phonological studies have achieved satisfactory descriptions of rhythm. This problem is mainly due to the traditional divide between phonology and phonetics: phonetics has not profited from recent phonological advances in the study of rhythm, while phonology has ignored for a long time the results of phonetic studies of rhythm.

Phoneticians, on the basis of impressionistic data, have traditionally seen rhythm as isochrony, i.e. as one or another type of events (stresses or syllables) occurring at regular intervals (e.g. Abercrombie 1967). This view gave rise to the traditional distinction between stress-timed and syllable-timed languages. The existence of the two categories remained virtually unquestioned[1] until the advent of more sophisticated instrumental techniques which enabled phoneticians to search for acoustic evidence for the two rhythmic categories. To their surprise they failed to find evidence for isochrony; neither interstress intervals in English, the stress-timed language *par excellence*, nor syllables in syllable-timed languages like French or Spanish, turned out to be of equal duration (see, for example, O'Connor 1965, Lea 1974 on English; Wenk & Wioland 1982 on French; Pointon 1980 on Spanish).

On the other hand, metrical phonology, the main phonological theory of rhythm, has had more success than phonetic theories by viewing rhythm as alternation rather than isochrony, and by representing rhythm in an efficient way through hierarchical structure (e.g. Liberman & Prince 1977). The main weakness of metrical phonology has been the adoption of principles which have proved suitable for the rhythmic representation of English for the representation of other languages. In particular, phonological studies of rhythm have often accepted uncritically the idea that the rhythmic structure of all languages, and consequently the phonological representation of this structure, is based on exclusively binary patterns. This assumption, however, has not yet been subjected to rigorous experimental testing. The relatively few acoustic studies which have tested phonological models do not find strong evidence in support of these models (e.g. Farnetani & Kori on rhythmic stress in Italian; for a discussion of this study see Chapter 5, section 5.3). Thus, it is necessary to validate phonological theories by testing them experimentally; in particular, it is necessary to test the assumption of binarity by studying languages which do not seem to have binary rhythm. One such language is Greek.

In this thesis I examine Greek both from a phonetic and a phonological perspective so that phonological hypotheses can be tested and refuted or improved on the basis of the experimental

[1] Except by Classe (1939).

results. Since the relationship between the perception and the production of rhythm is not yet clear, rhythm is studied both from the perception and the production point of view. The aim is to find which characteristics of the acoustic signal give rise to the perception native speakers have of their language's rhythm. One of the language's characteristics that is extensively examined in this study is stress, since it has often been assumed that rhythm is based on stress patterns. However, it is not taken for granted here that stress is the only contributor to rhythm. Rhythm is defined as the organisation of linguistic prominences, such as stress, full vowels and heavy syllables, pitch accents, etc. The exact nature of the prominences involved in the creation of rhythm is language specific and it is the interaction of these prominences that gives rise to a language's particular rhythm.

In the remainder of this chapter I examine, first, the phonetic evidence for stress and rhythm in speech, and second, the various phonological analyses of stress and rhythm. Finally, I present some essential features of the Greek phonological system and the phonological and phonetic analyses that have been put forward to account for some of the most problematic aspects of this system.

1.2 THE PHONETICS OF STRESS

1.2.1 INTRODUCTION

Prior to studying rhythm it is necessary to examine stress because of the close connection which is thought to exist between stress and rhythm[2] . For instance, Liberman and Prince (1977:250) state that "English stress patterns, within and among words, have phonetic reality as rhythmic patterns entirely independent of their role in orchestrating the placement of intonation contours". Although Liberman & Prince refer specifically to English, their view has implicitly been thought to represent the relation between stress and rhythm in other languages as well (e.g. Nespor & Vogel 1989 on Italian, Catalan and Greek). It is therefore desirable to see what are the acoustic and perceptual correlates of stress, since the results of the experiments presented here are interpreted in the light of such evidence.

1.2.2 PERCEPTUAL AND ACOUSTIC STUDIES OF STRESS

Stress has been related to force of utterance by earlier phoneticians. This idea can be traced at least back to Sweet (1906) and it was later adopted by Jones (1976[3]), Abercrombie (1967) and others (for extensive reviews of phonetic theories and research on stress see Lehiste 1970, Lea 1977, Ohala 1977 and Beckman 1986). The strongest advocate of this theory was Stetson (1951)

---

2 The connection between stress and rhythm has been disputed by scholars like Bolinger (1958) and Beckman (1986), though from different points of view. Bolinger claims that English has pitch accent, rather than stress accent, and thus denies the existence of stress as a prosodic phenomenon distinct from intonation. Beckman, on the other hand, disagrees with the theorists who equate stress (and consequently rhythm) with duration and argues for an integrated representation of accentual prominences relating both to stress and intonation (for a discussion, see Beckman 1986:62 ff.)

3 This date refers to the latest edition of Jones's *An Outline of English Phonetics*, first published in 1918.

who, on the basis of indirect data, put forward the idea that each syllable is pronounced with a *chest pulse* which is the result of increased activity of the respiratory muscles. According to Stetson, stressed syllables are related to stronger chest pulses which are produced with the additional participation of abdominal muscles. His theory was disproved by Ladefoged, Draper and Whitteridge (1958) whose electromyographic data showed that there is not a one-to-one correspondence between syllables and bursts of muscular activity in running speech. However, Ladefoged and his colleagues also reported that in many cases "there was an increase in the degree of muscular activity immediately before the syllables which were heard as being strongly stressed" (Ladefoged et al. 1958:211). Ladefoged et al's (1958) results are supported by Ladefoged (1963) and by Gay (1978), who cites Harris (1971) and Sussman & MacNeilage (1978) as finding that stress is accompanied by greater EMG activity of the relevant articulatory muscles.

However, the physiological correlates of stress are of less interest to the present study, which does not include physiological data, than its acoustic and perceptual correlates. What is of interest here is that stressed syllables were perceived as louder than unstressed syllables and that their loudness was attributed to force of utterance; relating stress to force of utterance meant that most phoneticians expected stress to have concrete physiological and acoustic correlates, namely high subglottal pressure and consequently high amplitude in the acoustic signal. Bloomfield (1935:110), for example, states that "*stress* - that is intensity or loudness - consists in greater amplitude of sound waves". These expectations, however, were not confirmed in their strong interpretation either by acoustic measurements or by perceptual experiments on stress.

In the nineteen fifties and sixties, most of the research on stress concentrated on the perceptual aspects of the phenomenon. In a classic series of experiments Fry (1955, 1958, 1965) tried to determine the relative importance of duration, amplitude, fundamental frequency and vowel formant structure as stress cues by manipulating two of the above correlates at a time while keeping the other two constant. He used the Haskins Pattern Playback to synthesise stimuli of the form *óbject* and *objéct* whose acoustic correlates were based on natural utterances. The results of Fry's three experiments show that the stress cues are, in order of importance, fundamental frequency, duration, amplitude and formant structure. The supremacy of F0 as a stress cue was corroborated by Bolinger (1958) and Morton & Jassem (1965).

The results of these studies indirectly contradicted the traditional idea that the main stress correlate is amplitude by showing that amplitude is not a robust perceptual cue for stress in English. However, it must be borne in mind that all these studies refer to stresses which carry the intonation nucleus. In this context the predominance of F0 as a stress cue is not surprising. The predominant role of F0 in stress perception is not supported by more recent research, which shows first that F0 is not as important a stress cue in non-nuclear position, and second that the role of amplitude and spectral structure in stress perception is not as insignificant as Fry's results imply.

Nakatani & Aston (1978) report on a perceptual experiment on stress, in which the influence of intonation was taken into account by embedding the test-words in sentences in which the test-words did not always occupy the nuclear position. Their results suggest that the importance of the various acoustic cues is related to the position of a word in the utterance. As expected, in nuclear position F0 was the overriding cue. In prenuclear position, duration and vowel quality were virtually as important as F0, while in postnuclear position duration became the most robust cue. Similar results are also reported by Huss (1978).

As mentioned, experimental work has also shown that amplitude and spectral structure are more important stress cues than Fry's studies suggest. Beckman (1986) found that F0 and amplitude were the most robust stress cues for test words in nuclear position, and that spectral structure was also an important cue. Beckman attributes the discrepancies between her results and Fry's partly to the fact that, due to the inherent limitations of the Haskins Pattern Playback used by Fry, amplitude and spectral information were rather unnatural in his stimuli and might not have been exploited by the listeners (for a discussion see Beckman 1986:195).

However, according to Beckman the different rating of amplitude in the two experiments is mainly due to the altogether different way of measuring and manipulating amplitude. In most experiments, including Fry's, peak amplitude is the value that is measured in natural utterances and manipulated in synthesised stimuli. Following these lines, Beckman indeed found that amplitude rated last as a stress cue, well below F0, duration and formant structure. Beckman, however, suggests that amplitude and duration information should not be separated but presented in a single measurement, which she calls *total amplitude* (the term used in this study is *amplitude integral*). This type of measurement ensures that when measuring amplitude the duration of a sound is also taken into account; thus a short sound with high average amplitude may have the same amplitude integral as a long sound with low average amplitude. This measurement is closer to the perceptual correlate of amplitude, loudness, which in short sounds, such as speech sounds, is significantly influenced by duration (Moore 1989:57 ff.). Beckman's results show that when amplitude is measured and manipulated in this way it becomes the most important acoustic and perceptual correlate of stress.

The importance of amplitude integral was first recognised by Lieberman (1960). Lieberman analysed acoustically 25 pairs of words of the type *cóntrast : contrást* embedded in meaningful sentences; he found that amplitude integral is the most robust stress correlate, while peak amplitude is a rather weak one. The significance of amplitude integral as a stress correlate is supported by the results of other studies. For instance, Lea (1977) reports on the results of Medress et al. (1971) and of Lea (1973). In both of these studies, the performance of computer programs devised for the location of stresses increased dramatically when the computer was instructed to look for amplitude integral rather than for duration or peak amplitude. For example, the program of Lea (1973) located successfully 84% of all stresses in connected speech using amplitude integral as a cue.

Experiments along the lines set by Fry have been performed for the investigation of stress perception in other languages as well. These experiments suggest that the importance of stress correlates is not the same in all languages. For example, Jassem, Morton and Steffen-Batóg (1968) used the material of Morton & Jassem (1965) with Polish subjects and found that duration was a more effective cue in Polish than in English. Similarly, Bertinetto (1980), who used synthesised versions of /'papa/ *pope* and /pa'pa/ *dad* for a perceptual test, concluded that Italian speakers rely on duration, amplitude and F0, in this order, in perceiving stress. Bertinetto's results are corroborated by those reported in Farnetani & Kori (1990).

Unfortunately, in some experiments certain prosodic features of the language under investigation were not taken into account, thus diminishing the reliability of the studies; one such case is Janota's (1967) study of Czech stress. Janota suggests that duration, amplitude and F0, in this order, are the stress cues in Czech. However, Czech has phonemically long and short vowels and dominant initial stress (Hyman 1977b). These characteristics of Czech are reflected in the fact that Janota's subjects showed a preference for the trochaic rather than the iambic pattern of his bisyllabic stimuli.

The above results show that any comparison between studies on stress in different languages should be cautious as a number of factors (like the ones mentioned for Czech) can interfere with the manipulation of acoustic correlates and bias the subjects' responses. Stress is relational in nature and its presence depends to a great extent on the linguistic system of which it forms part. This idea is shown convincingly by the following experiments.

Berinstein (1978) found that speakers of different languages are strongly influenced by the most common stress pattern of their language in their perception of stress. In her experiment only duration was manipulated in synthesised four-syllable words. The responses reflected the dominant stress-pattern of the subjects' native language: English speakers showed a preference for stress on the initial syllable, K'ekchi speakers for stress on the final syllable, and Spanish speakers showed no strong bias, since in Spanish words may carry stress in any of the last three syllables.

Similar results were found by Williams (1986). She asked two English and one Welsh subject to mark the stress of Welsh words in isolation and in short sentences elicited from one Welsh speaker. While the English subjects chose the final syllable of a word, the one that had longer duration and higher F0, the Welsh subject marked stress on the penultimate syllable (the most common stress pattern in Welsh), although this syllable showed consistently shorter duration and lower F0 than the final one. The results of Berinstein (1978) and of Williams (1986) strongly suggest that the importance of a listener's native language in stress judgements must not be underestimated.

1.2.3 CONCLUSION

To summarise, the correlates of stress in English are fundamental frequency, duration and intensity. Although perceptual studies suggest that the most important stress cue is F0 in nuclear position and duration in prenuclear and postnuclear position, acoustic studies show that amplitude is a more important stress correlate than F0 or duration. This discrepancy between acoustic and perceptual studies seems to be due to the different ways of measuring amplitude. If instead of measuring peak or average amplitude, amplitude integral is considered, amplitude becomes the strongest acoustic and perceptual correlate of English stress as Beckman's (1986) study demonstrates. Finally, experiments in languages other than English show that while the same parameters (i.e. F0, duration and amplitude) are used to signal stress, their relative importance is language specific.

1.3 PHONETIC STUDIES OF RHYTHM

1.3.1 EVIDENCE FOR RHYTHMIC CATEGORIES

As has been mentioned, phonologists have traditionally divided languages into stress- and syllable-timed ones. The terms were first used by Pike (1945) but the notion of two rhythmic categories had existed for much longer than that, as had the idea that English rhythm is organised by stresses which fall at regular intervals. Daniel Jones (1976:242), for example, says that in English "there is a general tendency to make the 'stress-points' of stressed syllables follow each other at equal intervals of time" and he also refers to the tendency of French to make continuous use of a rhythm in which all syllables are of equal duration. In stress-timed languages, isochrony between interstress intervals of varying length (measured in syllables) is supposed to be achieved by shortening the syllables of long feet, while lengthening those of short feet; in syllable-timed languages, isochrony between syllables of varying complexity is supposed to be achieved by temporal compensation within the syllable.

Classe (1939), the first to test experimentally the notion of stress-timing in English, concluded that there is some evidence for isochrony in English speech; however, for isochrony to exist, certain rather rare conditions, such as homogeneity of phonetic and syntactic structure of interstress intervals, must be fulfilled. Despite Classe's results and his caution, the notion that there are only two rhythmic categories persisted long after his work was published. Abercrombie (1967:97), for instance, asserts that "[a]s far as is known, every language in the world is spoken with one kind of rhythm or with the other."

The first studies that had some impact on the way phoneticians viewed rhythm were Shen & Peterson (1962), O'Connor (1965) and (1968), Bolinger (1965a), Uldall (1971) and (1972), and Lea (1974). (See Lehiste 1977 for a thorough review.) The results of all investigators agree that there is very slim evidence for isochrony of interstress intervals in English. All studies reported that foot length increased in relation to the number of syllables contained in the foot, although some studies reported greater variation than others. These results contradict the essence of stress-timing, which requires temporal compensation to take place in order to maintain constant foot length: no such compensation was found to take place.

Similarly, studies of languages other than English, particularly of languages which have been classified as syllable-timed, do not support the idea of two rhythmic categories. Neither French

(Wenk & Wioland 1982) nor Spanish (Pointon 1980), the syllable-timed languages *par excellence*, exhibit syllable isochrony. Some authors even suggest that Latin-American Spanish and Brazilian Portuguese (also classified as syllable-timed) share characteristics with stress-timed languages (Manrique & Signorini 1983 and Manrique & Massone 1984 on Spanish; Major 1981 on Brazilian Portuguese). However, the results of these studies are rather inconclusive, in that they provide evidence neither for syllable-timing nor for stress-timing: syllables exhibit durational variations, especially in relation to stress, while interstress intervals are not isochronous. Other authors refuse to classify the languages they study when they find that interstress intervals are not of equal duration, yet the length of syllables varies a great deal (Balasubramanian 1980 on Tamil; Toledo 1985 on Latin American Spanish).

The only language for whose rhythmic category there is supportive experimental evidence is Japanese, which is said to be mora-timed (Port, Dalby and O'Dell 1987; Hoequist 1983). Morae in most cases correspond to syllables, but syllables containing long vowels or geminate consonants count as two morae. Port, Dalby and O'Dell (1987) found that the duration of words increased linearly with the number of morae rather than syllables, although the durations of morae were not constant but depended on their segmental makeup, their position in the word and the word length measured in morae. The authors, however, admit that their remarkable results were obtained in ideal conditions which are not to be expected in natural (as opposed to laboratory) speech. Hoequist (1983) presents even stronger evidence for isochrony in Japanese by using reiterant speech, which might have caused the slight differences between his results and those of Port et al.

Various explanations have been put forward to account for the lack of experimental evidence for stress- and syllable-timing. For instance, Lehiste (1977) suggests that the great durational variation found among interstress intervals in English can be explained first, by the use of heterogeneous material in many studies and second, by the idea that isochrony is a perceptual phenomenon. However, there is experimental evidence against both Lehiste's hypotheses.

Concerning her first hypothesis, many studies have shown that considerable variation in foot duration cannot be eliminated by the use of more "homogeneous" material. For instance, Lehiste (1977) found that in sentences which contained only combinations of monosyllabic and bisyllabic feet, durational differences between the two foot types reached 133 ms or 26% of maximal foot length (500 ms), a perceptible difference (Lehiste 1977; see below). Most importantly, Nakatani et al. (1981) found no evidence for isochrony although they used reiterant speech, which eliminates the effects of segmental variation on rhythmic isochrony (see Liberman & Streeter 1978, and Larkey 1983 for validation studies of reiterant speech). Nakatani et al. showed that duration of feet and words increased linearly with their size (measured in syllables). Thus, even under seemingly ideal conditions, no evidence for isochrony was found.

For her second hypothesis, Lehiste (1977) draws on Allen (1975) who put forward the idea that isochrony and rhythm in speech are essentially perceptual phenomena. Lehiste concentrates on one of Allen's points, namely the psychological finding that humans tend to underestimate long intervals and overestimate short ones. This human tendency has been well documented by Woodrow (1951) and Fraisse (1963) with non-speech material like pure tones and noise bursts. Lehiste (1977), on the basis of her own research on speech, suggests that perceptual isochrony, although evident both in speech and non-speech stimuli, is more pronounced in speech.

As evidence supporting her hypothesis, Lehiste (1977) presents the results of an experiment in which her subjects heard utterances and non-speech material (noise intervals separated by clicks) with the same temporal pattern. The subjects were very good at distinguishing temporal differences in the non-speech material, while they perceived the speech stimuli as more isochronous than they were. Similar differences between speech and non-speech stimuli have been reported by Allen (1973) and Donovan and Darwin (1979). These results support Lehiste's hypothesis that perceptual isochrony is stronger in speech than in non-speech material. On the basis of such evidence, Lehiste reasons that the Just Noticiable Differences (JNDs) established by psychophysical experiments with non-speech stimuli might be shorter than those pertaining to speech, which she estimates at 10% of the foot duration for feet of 300-500 ms. As has been shown, though, differences greater than these JNDs have been reported by Lehiste herself concerning the length of monosyllabic and bisyllabic feet; so by this argument many *non-*isochronous feet should have been heard as such.

Moreover, Scott, Isard and de Boysson-Bardies (1985) did two experiments which suggest that perceptual isochrony is not related to rhythmic categories. The authors reasoned that interstress intervals should be perceived as more isochronous than they really are only by English subjects; French subjects and/or stimuli in French should not produce the same results. They tested both hypotheses: English and French subjects performed a tapping task both with English and French stimuli. The results showed that all subjects perceived both the English and the French stimuli as more isochronous than they were. The only difference between the French and the English subjects was that the French showed a greater tendency to perceive stimuli (especially the French ones) as isochronous! The authors hypothesised that the results were due to the complexity of the stimuli and to further test their hypothesis, they performed the same experiment using the English stimuli, noise bursts and a distorted version of the speech stimuli which was acoustically complex but not intelligible. While their subjects (English speakers only) showed no tendency for regularisation (i.e. for perceiving the stimuli as more isochronous than they were) as far as the noise bursts were concerned, they did regularise their tapping to the speech and distorted speech-stimuli. The authors concluded that perceptual isochrony as expressed by tapping experiments is not specific to speech but could be a mechanism used when the task becomes too difficult due to the complexity of the signal.

In my opinion, an alternative explanation could be that the subjects listened both to the speech and distorted-speech signals in a speech mode while they used a non-speech mode for the noise-bursts. This seems quite likely, since the distorted speech signals sounded rather like underwater speech and they could be identified as having once been speech (Isard, pers. comm.). If this explanation is correct, then perceptual isochrony could indeed be more pronounced in speech than in non-speech signals. One way to resolve the ambiguity would be to use acoustically complicated stimuli lacking speech qualities like harmonic structure; such stimuli would be as complicated as the distorted speech stimuli but would be more likely to be listened to in a non-speech mode. Scott, Isard and de Boysson-Bardies (1985) may not have a definite answer as to whether or not the human tendency for perceptual isochrony is stronger in speech and other complex signals or to speech only; however, their results with French and English subjects demonstrate convincingly that the perceptual isochrony of interstress intervals is not related to stress-timing.

Another theory based on the idea that rhythm is essentially related to perception is that of Dauer (1983, 1987). Dauer put forward the idea that languages do not fall into two distinct categories but occupy positions in a rhythmic continuum ranging from syllable- to stress-timing. A language's position on this continuum can be rated on the basis of certain phonetic and phonological criteria; for example, the relative duration of stressed and unstressed vowels, the variety of possible syllable structures in the language in question, the presence (or otherwise) of qualitative differences between stressed and unstressed vowels, etc. The "score" obtained by a language when these parameters are considered should indicate how close to one or the other category a language is: a language with more [-] signs should be closer to syllable-timing while a language with more [+] signs should be closer to stress-timing.

However, the system does not seem to give the results Dauer claims for it. For instance, Greek is given two [+] signs, one [0] sign and three [-] signs[4] for the parameters relevant to it. This, according to Dauer (1987:449), should mean that "[a]lthough the language may have some kind of accent, naive native speakers would have difficulty identifying the place of accent consistently in continuous speech, and linguists would have difficulty finding its acoustic correlates, even in words said in isolation." This is not true of Greek, however. In Dauer (1980a) there was very good agreement concerning the placement of accents in Greek running speech between phonetically naive native speakers of Greek and two trained phoneticians, one of whom did not speak the language. Furthermore, accent in Greek has a relatively high functional load (see 1.5 for details) and robust acoustic correlates (see Chapter 2).

Moreover, by attributing rhythm to the impression created by certain phonetic and phonological factors, Dauer implies that non-native speakers or at least trained phoneticians should be able to place a language on the rhythm continuum with some degree of consistency. On the contrary, Miller (1984) found that English and French phoneticians and non-phoneticians disagreed strongly on the rhythmic classification of eight languages (Arabic, Polish, Argentinian Spanish, Finnish, Japanese, Indonesian and Yoruba). One of the most striking results of his study was that Spanish, a language traditionally described as syllable-timed, was one of the two that

[4] The score is based on the "marks" given by Dauer for the categories for which she mentions Greek, and on my "marks" for the categories for which she does not. My "marks" are based on established facts about Greek prosodic structure (see 1.5).

were overwhelmingly identified as stress-timed[5] . Dauer's system gives Spanish the same score as Greek (the scoring is calculated in the same way as for Greek); this is clearly not the score one would expect from a language which is perceived as stress-timed.

In short, Dauer's theory offers good classification criteria for languages that can be easily classified, like English and French. However, her system cannot account correctly for the rhythm of languages like Greek and Spanish, which exhibit strong stress but not the alternation of stresses found in English. In this respect, the parameters she proposes are not sufficient for the classification of all languages on the rhythmic continuum she postulates. Moreover, her system of rating languages cannot be used for a formal representation of rhythm (cf. metrical phonology). Therefore, although the parameters mentioned in Dauer (1987) must be taken into account in studying a language's rhythm, their existence alone can neither explain how rhythm is created nor provide an abstract representation of linguistic rhythm.

1.3.2 CONCLUSION

In summary, a large number of studies in various languages have shown that the experimental evidence for rhythmic categories is at best very slim. Various proposals (Allen 1975, Lehiste 1977, Dauer 1987) have been put forward to account for this lack of evidence. Their common point is that rhythm is a perceptual phenomenon. However, there is experimental evidence (Miller 1984, Scott, Isard and Boysson-Bardies 1985) which questions the validity of this assumption. But, even if this assumption is correct, it neither provide us with the means of representing rhythm formally, nor explains how different languages are organised rhythmically. The results of experimental work suggest that if the linguistic study of rhythm is to provide a satisfactory analysis of the phenomenon three main points must be taken into account. The first is summarised very aptly by Roach (1982:78) who warns against dependence on the impressions of non-native speakers: "[...] the stress-timed/syllable-timed distinction seems at the present to depend mainly on the intuitions of speakers of various Germanic languages[6] all of which are said to be stress-timed; examination of the subjective feelings of speakers of languages usually classed as syllable-timed should be carefully studied if the distinction is to be maintained as a respectable part of phonetic theory."

The second point is that the study of rhythm should not be limited by unproved assumptions, for example that rhythm entails isochrony or that rhythm's only organising factor is stress. Unless more flexibility is incorporated into a working definition of rhythm (as suggested in section 1.1), experimental work will be hampered and limited by a priori assumptions. Finally, an abstract representation is necessary if the analysis of linguistic rhythm is to be complete.

---

5 The other was Yoruba which has been always considered a stress-timed language; see, for example, Abercrombie (1967:97).

6 It is perhaps significant that in Miller (1984) (see above) British phoneticians showed some agreement among them in classifying languages into rhythmic categories but their judgements rarely coincided with either those of non-phoneticians, or with those of French native speakers, trained or not.

1.4 STRESS AND RHYTHM IN PHONOLOGY

1.4.1 LINEAR MODELS OF STRESS SUBORDINATION

Although the results of various phonetic studies do not always agree on the acoustic and perceptual correlates of stress, the existence of stress has been virtually undisputed and its function in language described and analysed from various points of view.

The organisational function of stress and its part in the creation of speech rhythm have been recognised by most phonologists and phoneticians. Classe (1939:12), for example, states: "The phenomenon, which, by recurring at more or less regular intervals, creates what may be called, for want of a better phrase, a feeling of rhythm in speech, is generally admitted to be stress". The same idea is implicit in the term stress-timing and in the search for isochrony in interstress intervals; it is also present in more recent phonological analyses of rhythm. In Liberman & Prince's (1977) metrical theory abstract stress patterns are realised phonetically as rhythmic patterns. Similarly, Selkirk (1984) proposes that rhythm is based chiefly, though not exclusively, on stress.

Stress creates rhythmic patterns by subordination, i.e. by assuming different degrees of prominence. The idea that there exist various degrees of stress can be traced back to Sweet (1906:49) who claimed that the "degrees of stress are really infinite...", but suggested that in phonetic transcription it is sufficient to distinguish *strong, half-strong* and *weak stress*. Jones (1976) also distinguishes three types of syllables in relation to stress: *stressed syllables, syllables with secondary stress* and *unstressed syllables*. In the American tradition the different degrees of stress have been considered different phonemes. Thus, Bloomfield (1935) assigns three "secondary phonemes" to English: *highest stress, high* or *ordinary stress* and *low* or *secondary stress*. Trager and Smith (1951) distinguish four stress phonemes: *primary, secondary, tertiary* and *weak stress*.

Perhaps the most influential work on stress subordination has been *The Sound Pattern of English* (henceforth SPE) of Chomsky & Halle (1968). Chomsky & Halle address two questions: first, how to determine stress subordination within compound words and phrases, and second, how to assign word stress by rule. Their approach has three major characteristics. First, stress is considered a segmental property of vowels, i.e. a feature like [back] or [high]. Second, stress assignment is cyclic and depends on the lexical derivation of a word (for word stress) and on syntactic structure (for phrase stress). Third, stress is represented linearly, as numerals on top of the appropriate vowels: 1 represents primary stress, 0 lack of stress and all numbers above 1 represent diminishing degrees of stress. Stress subordination is implicit in the use of numerals to designate degrees of stress.

As a result of cyclical application of phrase-stress, each time a new word is added into a phrase, primary stress may remain in the same position but all other stresses are readjusted and reduced by one. For example,

```
 1  2          1  3        2        2        1  4        3
blackboard > blackboard eraser > John's blackboard eraser
   (1a)              (1b)                        (1c)
```

(SPE: 22).

There are two main problems with this practice. The first one is that as phrases grow bigger the notation changes the formal relationship between constituents: cf. *blackboard eraser* in (1b) and (1c). In other words, the SPE notation cannot adequately represent the fact that, in the above example, *blackboard eraser* is a unit within which the relationship of the constituents, namely that the stress of *eraser* is subordinate to the stress of *blackboard*, is fixed.

The second problem is that by demoting stresses every time a word is added to a sentence, the degrees of stress can reach absurdly high numbers. Such an infinite variety of degrees of stress, reminiscent of Sweet, is acoustically and perceptually implausible, however, despite Chomsky and Halle's (1968:25) assertion that "[...] a speaker who utilizes the principle of the transformational cycle and the Compound and Nuclear Stress Rules should 'hear' the stress contour of the utterance that he perceives and understands, whether or not it is physically present in any detail."

These and other problematic aspects of SPE have been addressed by a number of theories which sprang from the Chomsky & Halle generative background in the nineteen seventies and eighties (lexical phonology, metrical phonology, prosodic phonology, etc). The common point of these theories is that they all draw heavily from the work of Chomsky & Halle and that, in essence, they all try to devise more successful formalisms for the phenomena addressed in SPE, while retaining the basic underlying assumptions of generative phonology.

1.4.2 HIERARCHICAL MODELS OF STRESS

The most influential alternative to the SPE account of stress is metrical phonology, put forward by Liberman & Prince (1977) (henceforth LP). LP make the following suggestions: first, that stress is relational in nature, and second, that stress is not the property of vowels as the [stress] feature of SPE implies. LP suggest a different representation of stress to take the above points into account. Stress patterns are represented by hierarchical tree structures similar to syntactic trees; in fact metrical trees are based on syntactic constituency. In the metrical trees only binary branching is permitted; one of the two sister nodes is marked *strong* (s) while the other becomes automatically *weak* (w). The main point of the theory is that the labels *strong* and *weak* represent the relationship between the sister nodes. Consequently, the configurations [ss], [ww], [s] or [w] (alone) are meaningless. In this way, LP capture the relational nature of stress. The trees also capture the fact that prominence refers only to the relationship between sisters and does not influence any other nodes in the tree. For example,

```
(2a)        /\                          (b)       /\
           /  \                                  /  \
          /    \                                /    \
         w      s                              s      w
         |      |                              |      |
        warm   rain                          fruit   punch
```

The trees eliminate the need for cyclical application of phrase stress since the relative prominence of stresses is preserved under embedding: if (2b) is embedded in a longer phrase like (3) *a very refreshing fruit punch* its own structure as [sw] will remain the same within (3), an impossible option under SPE notation. Metrical theory also eliminates the need for numerous degrees of stress since stresses need not be demoted. By replacing linear with hierarchical structure, LP were the first to succeed in representing stress subordination in an economical and elegant way.

```
(3)

                   /\
                  /  s
                 /  / \
                /  /   \
               /  /     \
              /  /       \
             /  /         \
            /  /           \
           /  w             \
          /  /\              s
         /  /  \            /\
        /  /    \          /  \
       /  /      \        /    \
      w  s        w      s      w
      |  |        |      |      |
      a very refreshing fruit  punch
```

The metrical tree is used by LP for the representation of both phrase and word stress assignment. Word stress assignment is not of particular importance here, as LP's rules refer specifically to English. It is worth noting, however, that the use of the same formalism for phrase and word stress resulted in the assumption that binary stress patterns are the norm both within and across words. This assumption has been adopted by most subsequent theories of stress and rhythm. For example, Giegerich (1985) allows only trochaic feet in his model and postulates empty beats whenever the structure of words does not conform to a strictly trochaic pattern. Hayes (1981) distinguishes four types of metrical organisation, all based on binary branching trees.

1.4.3 N-ARY MODELS

Despite the general appeal of binary branching, there have been studies which rejected it in favour of n-ary tree structures. Given the extent to which phonologists examined stress assignment, it is not surprising that the seeds for n-ary models were sown by a theory of word stress assignment in English, Selkirk (1980).

According to Selkirk, word stress is not predictable by rule, as LP claim, but is lexical. Instead of using rules to determine the stress pattern of a word, Selkirk suggests that each word is linked in the lexicon with a specific tree template (among several possible templates) which determines the word's stress pattern. This tree template represents the word's *prosodic structure* in which the nodes are explicitly named, since the stress pattern of the word depends exclusively on them (the nodes). Thus, syllables (σ) are linked into stress feet (Σ) which in turn form phonological words (ω). The foot is the most important prosodic category for English word stress. For example, the difference between *raccoon*, which has two heavy syllables (in its American pronunciation), and *balloon*, which has only one, is formally incorporated in their different foot structures: *raccoon* has two monosyllabic stress feet, while *balloon* has one stress foot of iambic form. Formally, the two structures are as follows:

```
(4)   prosodic word        ω
                          /\
                         /  \
       foot            Σ_w   Σ_s
                        |     |
       syllable         σ     σ
                        rac coon
```

```
(5)   prosodic word        ω
                           |
       foot                Σ
                          /\
                         /  \
       syllable        σ_w   σ_s
                       bal  loon
```

In this way, Selkirk avoids the cumbersome solution proposed by LP to account for words like (4) and (5), which involves the lexical marking of syllables as [+stress] or [-stress] depending on their structure: heavy syllables (i.e. closed syllables and syllables with long vowels or diphthongs) are [+stress] and light syllables (i.e. open syllables with short vowels) are [-stress][7] .

Although in Selkirk's model monosyllabic feet become for the first time a legitimate part of the theory, rather than marked exceptions as in other metrical models, Selkirk's prosodic trees remain binary branching. However, the nature of the representation, i.e. the explicitly named nodes and the presence of monosyllabic feet as a matter of course, led other researchers to the conclusion that binary branching is superfluous in prosodic trees.

One such notable case is Nespor & Vogel (1986), who elaborated on the model first presented by Selkirk (1980). Nespor & Vogel increased the levels of representation to a maximum of seven: syllable (σ), foot (Σ), phonological word (ω), clitic group (C), phonological phrase (ϕ), intonational phrase (I) and phonological utterance (U)[8] . The justification for these levels is based on evidence from various prosodic phenomena, such as prominence patterns, and sandhi rules which take place within a domain but not across domain boundaries.

In the prosodic model presented by Nespor & Vogel (1986) the trees are n-ary branching. In addition, prosodic trees are constructed following rules specific to phonology instead of mapping syntactic constituency. Principally, prosodic trees must conform to the Strict Layer Hypothesis (henceforth SLH) which states that (a) trees cannot have recursive nodes and (b) only constituents of the same type can be sisters. The former constitutes the main difference between prosodic and syntactic trees which can have recursive nodes as, for example, in

(6) *The island that we visited last summer is in the Caribbean.*

In this sentence a S node (*that we visited last summer*) is under a NP node. In prosodic structure this is strictly prohibited: a phonological word node, for instance, cannot be subordinate to a foot node. The second requirement of the SLH results in a major difference between metrical and prosodic trees, namely that in metrical trees constituents of different levels, such as syllables and feet, can be sister nodes as in (7), while in prosodic trees they cannot.

(7)

```
          /\
         s  \
        /\   \
       /  \   \
      s   w   w
      ci  ne  ma
```

7 LP's use of the [stress] feature is an SPE vestige, although in LP [stress] does not play as important a part as in SPE.

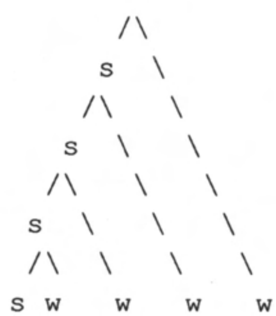
8 As Nespor & Vogel (1986) mention, a germinal form of their theory can be found in Selkirk (1981). Nespor & Vogel have added to Selkirk's model the Clitic Group.

Following the SLH, as presented by Nespor & Vogel (1986), (7) would have the prosodic structure in (8):

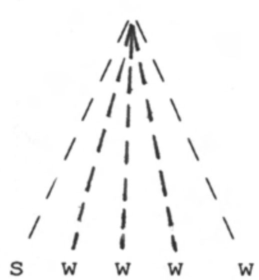
(8)

One argument in favour of n-ary branching is that it can account satisfactorily for the stress patterns of a variety of languages, some of which cannot be adequately represented by binary branching. Specifically, in languages like English binary tree structure is meaningful, in that it is related to stress placement and helps explain stress subordination[9] ; in languages with a single stress per word, however, the purpose of binary branching is obscure. Although hierarchical structure above the word is plausible in these languages as well, the rationale behind assigning a structure like (9) when (10) will do, remains unclear.

(9)



(10)



---

[9] However, research has shown that even languages like English can be represented by n-ary branching (Beckman 1986). Moreover, n-ary branching does not preclude binary branching.

The proponents of binary branching, posit (9) as the correct structure because one of the original points of LP was that only binary branching ensures the correct interpretation of trees (see section 1.4.2). However, according to metrical theory (see LP) trees can only show that *strong* is stronger than *weak*; i.e. binary structures, as presented by LP and their followers, can interpret relations between [s] constituents, and between [s] and [w] sisters but *not* between [w] constituents which are not sisters. In other words the structure in (9) does not contain more information than the structure in (10). Both structures show that [s] is stronger than all [w]s, but none makes any predictions about the relations between the successive [w]s. Therefore, the former structure is richer without presenting any advantage to compensate for its complexity.

Thus, if the convention that there is no relationship between [w]s is accepted, binary branching cannot be justified (for a more detailed criticism see Beckman 1986). If a different convention, which would allow for relations between [w]s, were accepted binary branching would proliferate the degrees of stress in the same way as the SPE model, which was abandoned precisely because it resulted in unrealistic distinctions between degrees of stress.

Moreover, Nespor & Vogel (1986) have pointed out that binary branching is not only superfluous for the representation of relative prominence between constituents, but also for the representation of other prosodic phenomena. Nespor & Vogel (1986) convincingly show that n-ary branching trees are as adequate as binary branching trees, without "making predictions that are not [...] borne out" and without adding a "constituent structure [which is] superfluous, [...] since there are no rules that make reference to them [sic]" (Nespor & Vogel, 1986:85).

Nevertheless, it must be borne in mind that, with the exception of Beckman (1986), the use of n-ary branching does not imply an acceptance of n-ary rhythmic structures. For example, concerning the foot level, Nespor & Vogel (1986:90) say that "it seems fairly unlikely that a language will in effect have only one strong syllable in a long polysyllabic word". In other words, although the formalisms may change from one theory to another, the underlying assumption remains that rhythmic patterns are binary.

1.4.4 THE PHONOLOGICAL REPRESENTATION OF RHYTHM

In addition to the tree, which represents abstract prominence patterns, LP postulate a distinct representation of rhythm, the metrical grid, which is based on the metrical tree and shows the timing of the events represented by it. (Thus, the grid implies that rhythm is based entirely on durational differences related to stress.) LP have both theoretical and practical reasons for proposing separate representations for stress and rhythm. First, according to them, the tree represents abstract prominence relations among constituents, while rhythm is simply the phonetic realisation of these abstract patterns. Therefore, the two cannot be represented by the same formalism.

Second, the grid is required if phenomena like stress clashes and the ensuing stress shifts, as in the much-cited example *thírteen mén*, are to be accounted for. According to LP, these phenomena can be captured much more easily by the grid than by the tree structure. LP explain the elimination of stress clashes by postulating that speech rhythm follows a *Rhythm Rule* whose purpose is to "create a (more nearly) alternating pattern, by eliminating perniciously close, or 'clashing', stresses" (LP:312).

In the grid each syllable is initially given one X; the rest of the grid is constructed in successive levels corresponding to nodes in the metrical tree, with greater stress being associated with more Xs. Clashes arise if at any given level successive Xs are not separated by at least one X at the lower level; they are remedied by moving the first one of the clashing Xs to an appropriate column on the left. Thus, the structure in (11a) will change to (11b) because at the third level of the grid the two Xs are not separated by an X at the second level.

```
(11a)                                    (b)

                x                                          x
          x_____x                        x                 x
     x    x     x                        x     x           x
     x  x x x   x                        x   x x x         x
     achromatic lens           >         achromatic lens
```

Most other theorists accept LP's view that abstract stress patterns and rhythmic structure must be represented by distinct formalisms, the former by metrical or prosodic trees and the latter by the grid (e.g. Nespor & Vogel 1989, Hayes 1984). However, this distinction has often come under criticism from different points of view. The most common criticism is that phonology does not need both the tree and the grid, since the grid maps exactly its corresponding tree structure. Studies have adopted both possible solutions to this problem of two representations. For example, Giegerich (1985) eliminates the grid as redundant and so does Beckman (1986) who presents a revised metrical tree for the representation of both stress and rhythm. On the other hand, Prince (1983) and Selkirk (1984) opt for a grid-only metrical representation.

The conflict between proponents of grid-only, tree-only or tree and grid representations is not of great consequence at this moment; it is discussed in more detail in Chapter 6, section 6.5.4. What is of importance is that in all models *eurhythmy* is assumed and rhythmic stresses, whose purpose is to create alternating rhythmic patterns, are postulated. For instance, Hayes (1984), on the basis of English and Polish examples, proposes the replacement of the Rhythm Rule of LP by "rules of eurythmy" which he believes to be (probably) universal. Similar concepts are expressed by Selkirk (1984) and Nespor & Vogel (1989), who suggest that alternating rhythmic patterns are universally preferred. What most scholars seem to have forgotten is one of the original remarks by Liberman & Prince (1977:250): "[s]ection 3 will introduce the concept of alignment with a *metrical grid*–fundamentally a formalization of the traditional idea of 'stress-timing' [emphasis added]." It seems that metrical phonology, by virtue of successfully representing stress subordination and rhythm in English, has been adopted by many scholars who automatically assumed the universality of the metrical principles originally postulated for English alone.

1.4.5 CONCLUSION

To summarise, the most important breakthrough in the study of stress and rhythm has been the replacement of linear by hierarchical representation as proposed by Liberman & Prince (1977). Various models whose purpose was to improve on the formalisms proposed by LP have been put forward, including models which are not based on binary, but on n-ary branching. However, despite the existence of diverse proposals on formalisms, the underlying assumption remains that rhythmic patterns are ideally binary. Unfortunately, the evidence for the universality of binary rhythmic patterns comes from a very limited number of languages. One language which does not seem to conform to the rules set by the phonological studies of rhythm is Greek, to which I now turn.

1.5 SOME ELEMENTS OF GREEK PHONOLOGY

Before examining the various phonetic and phonological studies of stress and rhythm in Greek, it is essential that certain undisputed features of Greek phonological structure are presented. First, in Greek, lexical stress conforms to a Stress Well Formedness Condition (henceforth SWFC), which allows lexical stress on any one of the last three syllables of a word but no further to the left (among others, Mirambel 1959; Malikouti-Drachman & Drachman 1980; Joseph & Warburton 1987). Because of the SWFC, when a suffix is added to a word with antepenultimate stress, lexical stress moves one syllable to the right of its original position; e.g.

(12) /a'martima/ *sin* > /a'martima+ta/ *sin+s* >/amar'timata/ *sins*[10] .

Second, as can be seen from example (12), lexical stress placement may depend on morphological factors, but it cannot be predicted from a word's metrical structure because (a) stress is not fixed and (b) there are no distinctions of phonological weight among the Greek vowels, /i, e, a, o, u/, or among syllables of different structure; i.e. in Greek, all syllables are of equal phonological weight (see, for instance Mirambel 1959:26-27, Joseph & Philippaki-Warburton 1987:251-254). Thus, it is quite common for phonemically identical words to be distinguished only by stress placement; e.g.

(13a) /'xo ros/ *space* : (13b) /xo 'ros/ *dance, noun.*

It is equally possible to find words like

(14) /'pli θos/ *crowd*

and

---

10 In the Greek examples the IPA notation of stress is followed with minor typographical alterations: ' designates primary stress and . secondary stress (Chapters 3 and 4) or rhythmic stress (Chapter 5) on the syllable following the symbol. Also, in phonological transcriptions /δ/ is used instead of /ð/, due to typographical limitations. Finally, all transcriptions are purely phonological and do not show allophonic variations, such as the palatalisation of velar obstruents, with one exception: the phonemic sequence /i/+V is presented as such if the two vowels belong to different syllables, e.g. /δu'lia/ *slavery*. If /i/ and the following vowel belong to one syllable, /i/ is represented as /j/, e.g. /δu'lja/ *work* . This is done in order to avoid confusion about the placement of stress in such cases.

(15) /'plin θos/ *brick*,

which are both stressed on their first syllable, although this is open in (14) and closed in (15).

Finally, the SWFC can also be violated by the addition of an enclitic to a host stressed on the antepenultimate syllable. If this SWFC violation takes place an *enclitic* stress is added two syllables to the right of the lexical stress (among others Warburton 1970; Botinis 1989). For example,

(16) /'maθima tu/ > /'maθi'ma tu/ *his lesson*

(17) /'δose mu to/ > /'δose 'mu to/ *give it to me.*

Notice that the type of change in the stress pattern of (16) and (17) is different from that in (12). Also, in (17) the second stress is added on the enclitic itself, since the host has only one syllable following its stressed one. This is the only case in which a clitic is stressed in Greek unemphatic speech. In emphatic speech articles can be stressed as in

(18) /'ine 'o poδosferi'stis/ lit. *He is THE football player*, met. *He is the best football player* (Setatos 1974).

These three points, namely the SWFC, the unpredictability of stress on the basis of a word's metrical structure, and the addition of a stress in a host-and-clitic group when the host is stressed on the antepenultimate are essential points from the perspective of this study. They are also the only points about Greek prosody on which there is (virtually) total agreement among researchers.

1.6 PHONETIC AND PHONOLOGICAL STUDIES OF GREEK STRESS AND RHYTHM

The past decade has seen the appearance of a number of studies of Greek prosody both in phonology and in phonetics. Studies in phonology include Malikouti-Drachman & Drachman (1980), Nespor (1988), Nespor & Vogel (1986) and (1989); studies in phonetics include Dauer (1980a), Fourakis (1986) and Botinis (1989). These are the first extensive studies concerning stress and rhythm in Greek. Various other studies deal in less detail with these two phenomena (Mirambel 1949, 1959; Warburton 1970; Sotiropoulos 1972; Setatos 1974; Theofanopoulou-Kontou 1978; Magoulas 1979; Joseph & Philippaki-Warburton 1987).

All these studies agree on the three points mentioned in section 1.5, but show remarkable disagreement on other aspects of Greek phonology. One point of disagreement concerns enclitic stress. As mentioned in section 1.5, a host-and-clitic group with stress on the host's antepenult will acquire a second stress. Almost all scholars who mention the phenomenon accept that this is the case[11] and that the two stresses in the host-and-clitic group have different prominence values (Warburton 1970; Setatos 1974; Malikouti-Drachman & Drachman 1980; Joseph & Warburton 1987; Botinis 1989). However, they disagree on the relative prominence of the two stresses. Most scholars agree that the added stress is stronger than the host's lexical stress (Warburton 1970;

[11] The only exceptions are Mirambel (1959) and Magoulas (1979),who claim that the enclitic stress may remain the only stress in a host-and-clitic group. This is a rather radical position which disagrees entirely with the present experimental data, the opinion of the majority of scholars, and the intuitions of native speakers. For these reasons this position will not be discussed further.

Malikouti-Drachman & Drachman 1980; Dauer 1980a; Joseph & Philippaki-Warburton 1987). Joseph and Philippaki-Warburton (1987:243), for instance, state that "[o]f the two stresses the last one is stronger, thus primary, while the original one is weakened to secondary". Botinis (1989) presents an entirely different analysis. Influenced by work on Swedish prosody, he claims that the SWFC-induced stress is a *phrase stress*, i.e. acoustically and perceptually distinct from *word stress*. Botinis's rather problematic analysis will be discussed in detail in Chapter 3, section 3.4. Suffice it to say here that Botinis's empirical evidence suggests that the SWFC-induced stress is acoustically the most prominent in the host-and-clitic group. Setatos (1974) however, followed by Nespor & Vogel (1986, 1989), claims that the host's lexical stress remains the strongest. Thus, a first point of disagreement emerges, namely the prominence value of the enclitic stress.

A second point of disagreement is the presence of rhythmic stresses in Greek. Most of the studies mention stress subordination only in relation to the host-and-clitic group stress addition, in which case they usually refer to *secondary stress* (Warburton 1970; Dauer 1980a; Botinis 1989). In other words, in most of the studies it is assumed that in Greek each word normally carries only one lexical stress. In many cases this is stated quite explicitly. For instance, Warburton (1970:37) says that "every non-monosyllabic word has one stress and only one". Sotiropoulos (1972:27), obviously influenced by American structuralism, states that "Modern Greek has one stress phoneme". Botinis (1989:15) proposes the *Monotonic Principle*, according to which "[e]very word that belongs to a major part of speech, i.e. noun, adjective, verb, adverb, has only one word stress [...]".

In contrast with the above studies, some phonological analyses assume that, in addition to lexical stress, Greek exhibits rhythmic stresses which are added at the surface level to maintain an even rhythm (Joseph & Philippaki-Warburton 1987; Malikouti-Drachman & Drachman 1980; Nespor & Vogel 1989). To avoid confusion, it is necessary to give a definition of the various attributes of the word *stress* that are used in this study in relation to Greek. The terms *lexical* and *primary stress* always refer to the main stress of a word. The term *secondary stress* refers to the host's original stress in a host-and-clitic group which violates the SWFC. The terms *enclitic stress* and *SWFC-induced stress* refer to the stress added in such cases. Finally, the term *rhythmic stress* refers to all other stresses (except lexical, secondary and enclitic) which are said to be added in order to preserve the even rhythm of an utterance. This usage is constant throughout the thesis unless the terms appear in a report of another study in which they have a different meaning; in such cases these terms are in quotes.

Of the scholars who postulate that Greek exhibits rhythmic stress, Joseph & Philippaki-Warburton (1987:243) take a very moderate view, suggesting that "[t]here are occasionally two stresses on a single grammatical word. This however is optional and it applies only to long words where the stressed syllable is preceded by several unstressed syllables, e.g. [proγramati'kos] [*programmatic*] with the only basic stress on the ultima or [.proγramati'kos] with the basic stress remaining primary and another (secondary) on the first syllable."

Nespor & Vogel (1989) (henceforth NV[12]) and Malikouti-Drachman & Drachman (henceforth M-DD) take a much stronger approach to rhythmic stress. They postulate that in Greek rhythmic stresses are added every time there is a succession of more than two unstressed syllables. In addition, NV (1989) and M-DD (like Joseph & Philippaki-Warburton) relate rhythmic stress to the secondary stress of host-and-clitic groups but in two different ways.

NV (1989:70) propose that rhythm is represented by the grid which is built "on the basis of the prosodic structure of a given string", and which helps resolve rhythmic anomalies such as stress clashes and lapses. NV (1989) suggest that in Greek rhythmic stresses appear when there is a lapse in the first level of the grid, i.e. a series of more than two unstressed syllables. When a lapse occurs, one of the unstressed syllables acquires a second asterisk in the grid[13] , i.e. a rhythmic stress, through the *Beat Addition* rule. As has been mentioned, NV (1986, 1989) also maintain that the enclitic stress (or "secondary stress" as they call it) is less prominent than the original lexical stress of the host; since the main stress on a ω is represented by three asterisks in the grid, it follows that syllables with "secondary stress" can only have two asterisks. In other words, rhythmic stress and secondary stress have the same phonetic realisation, since (a) they are both represented by two asterisks in the grid and (b) the grid cannot show constituency differences. For example, the sentences

(19) [to 'prosopo]$_C$ [mu 'itan ɣno'sto]$_C$ *the face was familiar to me*

and

(20) [to 'proso.po mu]$_C$ ['itan ɣnosto]$_C$[14] *my face was known*

have the same grid, displayed in (21); /po/, the only syllable with two asterisks, has rhythmic stress in (19) and enclitic stress in (20).

(21)

```
                        *
                        *
     *                  *
     *        *         *
     *   *    *         *
 *   * * *  * * *    *  *
to prosopo mu itan ɣnosto
```

According to NV (1989) then, the only difference between an enclitic and a rhythmic stress is that the former is the result of an obligatory prosodic rule which operates within the Clitic Group (C), while the latter is the result of *Beat Addition*, an optional rhythmic rule which operates in the grid; this difference, however, is not perceptible in speech.

---

12 The acronym NV refers to both Nespor & Vogel (1986) and to Nespor & Vogel (1989) unless otherwise stated.

13 NV (1989) use asterisks rather than Xs for the grid.

14 The stresses on /'prosopo/ are presented here with the prominence values assumed by NV (1989).

As mentioned, M-DD also relate secondary and rhythmic stresses, but their approach differs from that of NV, in that M-DD assume that the secondary stress in a host-and-clitic group is the stress of the host. Rhythmic stresses are added following the *Rhythm Rule* which states: "[m]ake a trochaic foot of *any* adjacent pair of weak syllables to the left of the lexical stress within the word [word+clitics] (iterative)" (M-DD 1980: 284). For example,

(22)

```
                        /\
                       /  s
                      /  /\
      /\             /  /  \
     / s            /  w    s
|   / /\    |      /  /\    /\
w w  w s    w      w s  w s  w
i ekloγi mu    >   i ekloγi mu
```

my election

As can be seen from (22) and (23), the Rhythm Rule applies not only to the left but also to the right of the lexical stress; this "refooting" explains the SWFC-induced stress. However, by equating the metrical structures of (22) and (23), this analysis cannot differentiate between a metrical tree with an optional rhythm-induced stress like (22), and a tree with an obligatory SWFC-induced one like (23).

(23)

```
                          /\
                         /  s
                        /  /\
     /\                /  /  \
     s \              /  w    s
|   /\  \   |        /  /\    /\
w s w   w   w        w s w  s  w
i epavli mu    >     i epavli mu
```

my villa

In other words, while both NV and M-DD give the same metrical representation to rhythmic stress and the less prominent stress of a host-and-clitic group, NV equate rhythmic stress to enclitic stress, and M-DD equate rhythmic stress to secondary stress (i.e. the host's stress).

To summarise, there seem to be two main interconnected issues that are addressed by studies of Greek prosody, namely the presence and nature of enclitic and rhythmic stress in Greek. In brief, NV and M-DD claim that Greek exhibits rhythmic stress; although they disagree as to which of the two stresses in a host-and-clitic group is the most prominent, they agree that the weaker one of the two is identical to rhythmic stress. Botinis (1989), on the other hand, does not mention rhythmic stresses but proposes two distinct prosodic categories, word and phrase stress, to account for enclitic stress. These issues are of great importance from the point of view of the present study, as they are concern crucial aspects of Greek rhythm, such as stress subordination and stress patterns. These issues are addressed here by means of acoustic and perceptual tests rather than impressionistic data.

1.7 AN OVERVIEW

Before proceeding with the examination of rhythmic aspects of Greek, in Chapter 2 I examine the acoustic correlates of Greek primary stress in a simple acoustic environment. This step is taken for two reasons: since stress correlates differ somewhat among languages, as has been shown, it is necessary to establish what these are for Greek; this is an essential step, since most of the subsequent experimental work examines stress. In addition, the data presented in Chapter 2 provide evidence for certain rhythmic aspects of Greek, in particular the elimination of stress clashes and word-final lengthening. After establishing the acoustic correlates of stress, it is possible to address the following questions: (a) whether the enclitic stress is the most prominent in the host-and-clitic group, as NV say; (b) whether the enclitic stress is perceptually distinct from a lexical stress, as Botinis suggests. These two issues are examined in Chapter 3 by means of a perceptual experiment and acoustic analyses of natural speech. The same means are used in Chapter 4 to address a third question, namely whether or not secondary stress and rhythmic stress are perceptually and acoustically the same, as M-DD suggest. Chapter 5 reports on further research on the acoustic correlates of rhythmic stress and also examines phonological evidence for rhythmic stress. Finally, Chapter 6 is a summary and discussion of the issues addressed in the previous chapters, with emphasis on the implications of the results for the rhythmic structure of Greek. Three main issues are addressed: the acoustic and phonological evidence for the elimination of clashes and lapses in Greek, the question of an appropriate formalisation of the results within phonological theory, and the consequences of the Greek data for the phonology and phonetics of rhythm.