

3. THE PERCEPTION AND ACOUSTICS OF ENCLITIC STRESS

The purpose of the perceptual test and acoustic analyses presented in this chapter is to examine the nature of enclitic stress. It has been argued by Setatos (1974) and Nespor & Vogel (NV) (1986) and (1989) that in a host-and-clitic group with antepenultimate stress, the stress that is added in accordance with the Stress Well Formedness Condition (SWFC) is less prominent than the lexical stress of the host, while other studies take the opposite view (e.g. Joseph & Warburton 1987; Dauer 1980a). A solution of this controversy is necessary, as the relative prominence of the two stresses in a host-and-clitic group has important implications for the phonological description of the phenomenon (see section 3.4 and Chapter 6, sections 6.5.3) and also for the relation of secondary with rhythmic stress (see Chapter 4). The empirical data presented here aim to solve this dispute; they are also used to examine whether Botinis's phrase stress and word stress are two distinct categories as his analysis suggests (see Chapter 1, section 1.6). The perceptual test and the acoustic analyses will be referred to as Experiment 3.

3.1 EXPERIMENT 3: METHOD

3.1.1 MATERIAL

For the perceptual test and the acoustic analyses two test pairs were designed (see the underlined parts of Table 1). In each test pair the two members are phonemically identical but have word boundaries at different places and are orthographically distinct in Greek. The first member of each pair, (a), consists of two words which together form a phrase and which are stressed on the same syllables as member (b). The second member, (b), consists of one word stressed on the antepenultimate and followed by an enclitic possessive pronoun. As this pattern violates the SWFC, a stress is added on the last syllable of the host. Thus, the difference between members (a) and (b) of each test pair is that in (a) each one of the two words carries lexical stress, whereas in (b) the phrase contains a lexical and a SWFC-induced stress on the same syllables as (a). According to NV (1986, 1989) the most prominent stress in (b) phrases is the lexical stress of the host, while in (a) phrases it is the stress of the second word (i.e. the one that falls on the same syllable as the enclitic stress in (b)) since the second word is the head of the phonological phrase, ϕ (NV 1986:165 ff.). Also, in Botinis's terms (a) and (b) phrases have different stress patterns, (a) containing two word stresses and (b) containing one word and one phrase stress¹; these stress patterns are said by Botinis to be perceptually distinct. If either NV or Botinis are correct, (a) and (b) phrases should be distinguishable.

¹ It could be argued that both (2a) and (2b) carry sentence stress, the prosodic category above phrase stress in Botinis's system. Even if this is the case with test pair (2), the distinction between word and phrase stress in (1a) and (1b) remains intact.

Table 1:

The test phrases, underlined, in the context in which they were read.

1

(a) /e'ɣo tu fo'nazo 'ari 'stasu kja'ftos ðe stama'tai/

I shout at him "Ari stop" but he doesn't [stop].

(b) /tu 'ipa ja to .ari'sta su ke 'xarike po'li/

I told him about your 1st class mark and he was very pleased.

2

(a) /'ixan a'nekaθen psa'ra ði'ko tus/

They have always had their own fishmonger.

(b) /pso'nizi 'panda a'po to psa.raði'ko tus/

S/he always shops from their fishmongery.

The test phrases were incorporated into meaningful sentences (see Table 1). Care was taken to avoid stress clashes, and to design, for each pair, sentences of similar prosodic structure and length. Due to the difficulty of finding pairs of test phrases fulfilling the stress requirements, it was not possible to control for other factors, such as ease of segmentation; for instance, in (1) the test phrase begins with a vowel and is preceded by another vowel. This could not be avoided because Greek grammar requires each noun to be preceded by an article. No other words could have been used instead.

Two distractor pairs were devised to be superficially similar to the test pairs and were also incorporated into meaningful sentences (see Table 2). The difference between test phrases and distractors is that in the distractors one member contains two words, each one with its own lexical stress (e.g. /'mono 'loɣo/ *only reason*), while in the other member the same sequence of syllables makes up one word with lexical stress on a different syllable from those stressed in the first member (e.g. /mo'noloɣo/ *monologue*).

3.1.2 SUBJECTS

Four speakers took part in the recording including myself. Two of the female speakers, HP and AA, and the male speaker, AP, were postgraduate students at the University of Cambridge, and were in their twenties. The fourth speaker, KAP, was a 60 year old woman visiting Cambridge. All were native speakers of Greek and spoke the standard dialect. All, apart from KAP, had extensive knowledge of English. None of the speakers had any history of speech or hearing problems. Apart from the author all speakers were naive as to the purpose of the experiment.

Table 2:

The distractor words, underlined, in the context in which they were read.

1

(a) /pi'stevo 'oti 'ksero to 'mono 'loyo ja a'fti tin ka'tastasi/
I believe that I know the only reason for this situation.

(b) /ðen 'exo a'kusi 'pjo vare'to mo'noloyo sto 'θeatro/
I haven't listened to a more boring theatrical monologue.

2

(a) /ðe 'θelo 'pare 'ðose me a'fto to 'atomo/
I don't want to have anything to do with this person.

(b) /'ksero 'oti to pa'reðose stus ðike'uxus/
I know that s/he delivered it to the beneficiaries.

Eighteen native speakers of Greek (7 male and 11 female) did the perceptual test in Athens, Greece. All listeners spoke standard Greek, as spoken in Athens, where sixteen of them live; the other two listeners live in the south of Greece and at the time of the experiment were visiting Athens. Twelve of the listeners were between 25 and 40 years old, and the other six were between 40 and 60 years old. Fourteen of them spoke other languages in addition to Greek, but only one had extensive knowledge and contact with a foreign language (English). All listeners had at least secondary education and fourteen of them held University degrees. None of the listeners had a history of speech or hearing problems. They were all naive as to the purposes of the experiment.

3.2 PERCEPTUAL TEST: PROCEDURE AND RESULTS

3.2.1 PROCEDURE

The recording took place in a sound treated room in the Phonetics Laboratory of the Linguistics Department, Cambridge University. Each of the four speakers read the test sentences and the distractors six times from a randomised list, typed in Greek. Prior to the recording the speakers were given some time to familiarise themselves with the material. They were also instructed to read the sentences as naturally as possible and to avoid using contrastive stress. In the case of test sentence (1a), they were also told not to read the test phrase, /'ari 'stasu/ *Ari stop*, as if it were in quotes; instead, they were told to read the sentence without pausing either between the carrier and test phrase or between the two words in the test phrase. During the recording the performance of the speakers was monitored, and if they did not follow the instructions they were asked to repeat the sentence. This was done in both the test sentences and

the distractors so as not to hint that only some of the material mattered.

The recorded sentences and the distractors were low-pass filtered at 7.8 kHz² and digitised at 16 kHz. For each test phrase and distractor one token from each one of the four speakers was selected for the test tape. The tokens chosen were those that sounded most natural to me. The main aim was to avoid coarticulatory interference from the carrier phrase as much as possible. This was important because the selected sentences for the test tape were edited so that only the test and distractor phrases were left. If there had been different coarticulatory influences at the boundaries of different stimuli, the listeners' responses might have been biased by them. In addition, I chose tokens which sounded, as much as possible, as if they had been spoken in isolation, since I did not want the listeners to be distracted by the fact that the stimuli were excised from longer sentences.

The test phrases and the distractors were recorded on cassette tape using computer-generated randomisation by blocks so that each token from each speaker was heard twice. Each test phrase and distractor was preceded by a 50 dB warning tone. There were 100 ms of silence between the tone and the following phrase and 2 sec between each stimulus and the following tone. Every 20 stimuli there was a 5 sec pause. In order for listeners to familiarise themselves with the task, the first 4 stimuli were repeated at the end of the tape, and the first 4 responses of each listener were discarded. Thus, each listener heard a total of 70 stimuli: 4 speakers x (4 test phrases + 4 distractors) x 2 blocks + 4 repeated items + 2 stimuli that consisted of two tones each (a result of the randomisation program).

The listeners did the perceptual test using a portable Sony Stereo Cassette-Corder TCS-450 (a good quality "walkman") and its headphones. Each listener did the test on his/her own. Half of the listeners did the test in my house or in theirs, in a quiet room where they were left alone. The other half did the test in their place of work; half of these listeners could retire to a quiet room to do the test, while the other half did it in a room where there were other people present. No subject complained that their performance might have been marred by noise or poor quality equipment, and no differences were observed among the responses of listeners who did the test under different noise conditions.

The listeners were given a response sheet, typed in Greek, which gave both possible interpretations of every stimulus in the tape (70x2 possible answers). It was explained to them that the phrases they were going to hear were excised from longer stretches of speech and were preceded in the tape by a warning tone; they were also told that their task was to choose which of the two interpretations on the answer sheet corresponded to the phrase they heard. The listeners were urged to give an answer to all stimuli, even if they were not absolutely certain of their choice; they were not allowed to play back the tape.

² This concerns only the data used for the perceptual test tape, which was prepared from Audlab files. The data were digitised again at the same sampling frequency for the acoustic measurements for which ILS files were used; in this case the material was low-pass filtered at 8 kHz.

3.2.2 RESULTS

In the 576 responses to the test phrases (18 listeners x 32 test phrases/answer sheet) there were 290 mistakes. That is, half of the responses to the test phrases were wrong (see Table 3³). While all listeners made errors, the number of errors ranged from a minimum of 9 (one listener) to a maximum of 21 (one listener). In contrast, only two listeners made mistakes (1 and 4 respectively) in the distractors, whose identification rate was 99.1%.

Table 3:

Contingency table of type of stimulus by listener response.

(a) Observed responses

STIMULUS	RESPONSE		Total
	"one-word"	"two-word"	
"one-word"	113	175	288
"two-word"	115	173	288
Total	228	348	576

(b) Expected responses (and deviances)

	"one-word"	"two-word"
"one-word"	114 (0.008)	174 (0.005)
"two-word"	114 (0.008)	174 (0.005)

Total deviance (χ^2) = 0.026 1df

The difference between the relevant conditions is not significant.

3.3 ACOUSTIC ANALYSES: MEASUREMENTS AND RESULTS

3.3.1 MEASUREMENTS

For the acoustic analysis, duration, amplitude and F0 were measured. As in Experiment 2, ILS was used to measure F0, and Audlab was used to measure duration and amplitude. F0 and amplitude measurements were performed in the way described in Chapter 2, section 2.3.4, with the difference that peak amplitude was not measured for these data.

Due to the greater complexity of the material, durational measurements were obtained from spectrograms (rather than waveforms) using the Audlab facilities and following standard criteria of segmentation (see Peterson & Lehiste 1960). The most difficult point was measuring the initial /a/ of the test phrases (1a) and (1b), /fo'nazo 'ari 'stasu/ and /to .ari'sta su/ respectively. The

³ Table 3 shows first the number of one- and two-word responses by type of stimulus and second how the same number of one- and two-word responses would be expected to break down if the type of stimulus did not influence the listeners' choice. As can be seen from Table 3, two-word stimuli were correctly identified more often than one-word stimuli (173/288 vs 113/288 correct responses respectively). This difference may suggest that Greek native speakers prefer assigning one stress to each word to assigning two stresses (a lexical and an enclitic one) to the same word; using this strategy obviously results in more two-word stimuli being correctly identified.

point of maximum change in the value of the formants, especially F2, between /o/ and /a/ was used as the criterion for the segmentation. The accuracy of these measurements is within 3 pitch periods (approximately 12-18 ms for the female speakers and 24 ms for the male speaker); in the rest of the measurements the error range is one pitch period. The frication and silence of the /st/ sequences in /'ari'sta su/ and /'ari 'stasu/ were measured as a single unit. VOT was measured as part of the following vowel.

Measurements were obtained for the first three syllables of test phrases (1a) and (1b), /arista/, and for the three middle syllables of test phrases (2a) and (2b), /raðiko/. Only three tokens from each speaker were analysed. These were the first three tokens of each speaker, if in these was included the phrase used in the perceptual test; if not, the third token was replaced by the token used in the listening test. For convenience, the (a) member of each test pair will be referred to as the two-word member and the (b) member as the one-word member. The terms do not reflect the number of words in each test phrase but, rather, the number of words the three syllables under investigation belong to.

For each of the three syllables examined here the two members of each test pair are compared. The most important comparison concerns the last syllables, /sta/ and /ko/, which in the two-word members carry primary stress and in the one-word members carry enclitic stress (which is thought by some to be a secondary stress). The first syllables, /a/ and /ra/, are also compared in order to see whether the weakened lexical stress of the host in the one-word member is acoustically different from the lexical stress of the first word in the two-word member. Finally, the middle syllables, /ri/ and /ði/, are compared, in order to examine possible acoustic differences between them which would justify different tree structures for the one-word and the two-word versions, as discussed in section 3.4.

3.3.2 DURATION

Durational data for all speakers are shown in Table 4. The data were analysed using 2-way repeated-measures analyses of variance (word makeup x speaker), one for every pair of syllables. The data from each test pair were analysed separately because the segmental makeup of the relevant syllables differed substantially between the test pairs. Analyses of variance were also performed on the vowel durations, but vowel data will be discussed only if they differ from syllable data. The differences among speakers were examined in the few cases when there was an interaction between speakers and word makeup.

As shown in Table 5, there are no durational differences either between initial syllables or between middle syllables within each test pair. However, the vowel durations of the middle syllable /ri/ show an interaction between speakers and word makeup; this interaction is due to HP, in whose data the vowel duration of the one-word /ri/ is longer than that of the two-word /ri/ ($F(1,8)=6.9$, $p<0.028$).

Table 5 also shows that there are no durational differences between the syllables and vowels of one-word and two-word /ko/. The /sta/ syllable durations, however, show an interaction

Table 4:

Mean durations (ms) and standard deviations of the consonants, vowels and syllables of the two-word member (with #) and of the one-word member of /arista/ (top) and /raðiko/ (bottom), for all speakers.

		<u>consonant(s)</u>	<u>vowel</u>	<u>syllable</u>
a	Mean	-	-	117
	S.D.	-	-	17
ri	Mean	27	50	77
	S.D.	5	10	11
# sta	Mean	123	173	296
	S.D.	17	20	36
<hr/>				
a	Mean	-	-	119
	S.D.	-	-	21
ri	Mean	25	52	77
	S.D.	5	15	14
sta	Mean	115	168	283
	S.D.	12	22	28
<hr/>				
ra	Mean	27	104	131
	S.D.	5	12	13
# ði	Mean	48	37	85
	S.D.	7	7	10
ko	Mean	66	160	226
	S.D.	21	31	36
<hr/>				
ra	Mean	27	107	134
	S.D.	6	13	16
ði	Mean	46	38	84
	S.D.	9	9	15
ko	Mean	72	154	226
	S.D.	12	32	35

between speakers and word makeup, which is again due to HP: in her data the two-word /sta/ is longer than the one-word /sta/ ($F(1,8)=8.85$, $p<0.017$). As can be seen in Table 5, there are no durational differences between the vowels of the two /sta/s.

On the whole the durational results confirm those of the perceptual experiment by showing no major differences between the one- and two-word versions of the test phrases.

Table 5:

F-ratios and probability levels from 2-way ANOVAs for comparison of syllable duration (left) and vowel duration (right) in two-word and one-word test phrases. A * indicates speaker and word makeup interaction.

	<u>Syllable durations</u>		<u>Vowel durations</u>	
	<u>F(1,8)</u>	<u>p</u>	<u>F(1,8)</u>	<u>p</u>
a	0.21	n.s.	-	-
ri	0.21	n.s.	0.62*	n.s.
sta	6.17*	0.03	0.76	n.s.
ra	0.82	n.s.	0.97	n.s.
ði	0.36	n.s.	0.10	n.s.
ko	0.01	n.s.	1.09	n.s.

3.3.3 AMPLITUDE

The AI and RMS data for all speakers are shown in Table 6. AI means for each speaker are shown in Figures 1 and 2. The same procedure as for duration was followed for the statistical analysis of amplitude.

Table 7 shows that the AI of the initial syllables of /ari'sta su/ and /'ari 'stasu/ is the same. The AI of the initial syllables of one- and two-word /raðiko/, however, shows an interaction between speakers and word makeup. This interaction is due to KAP, in whose speech the one-word /ra/ has higher AI than the two-word /ra/ ($F(1,8)=7.49$, $p<0.024$). KAP's data are responsible for /ra/'s statistically significant F-ratio shown in Table 7; in the data of the other speakers there is no difference between the two /ra/s, as Figure 2 shows (HP: ($F(1,8)=0.38$, n.s.; AP: $F(1,8)=0.19$, n.s.; AA: $F(1,8)=0.63$, n.s.).

Table 6:

Means and standard deviations of the normalised AI and RMS of the two-word member (with #) and of the one-word member of /arista/ (left) and /raðiko/ (right), for all speakers.

		<u>AI</u>	<u>RMS</u>	<u>AI</u>	<u>RMS</u>	
a	Mean	69	163	58	175	ra
	S.D.	7	15	10	22	
ri	Mean	30	105	22	110	# ði
	S.D.	7	13	7	29	
# sta	Mean	57	110	47	117	ko
	S.D.	6	12	6	11	
<hr/>						
a	Mean	68	158	62	185	ra
	S.D.	8	14	7	18	
ri	Mean	27	95	23	116	ði
	S.D.	13	33	4	20	
sta	Mean	59	114	49	121	ko
	S.D.	5	12	7	16	

The AI of the one-word and two-word /ði/s is the same, while the /ri/ data show that some of the speakers differentiate between one- and two-word versions (see Figure 1). In particular, while AP and AA have the same AI in the two conditions (AP: $F(1,8)=0.62$, n.s.; AA: $F(1,8)=2.2$, n.s.), HP has significantly higher AI on the two-word /ri/ ($F(1,8)=10.49$, $p<0.01$), and KAP significantly higher AI on the one-word /ri/ ($F(1,8)=5.18$, $p<0.05$). There is no significant effect of word makeup on the AI of either of the final syllable pairs, /sta/ and /ko/.

The RMS data agree entirely with the AI ones. In the two comparisons which show speaker and word makeup interaction in the RMS data (see Table 7), the interaction is due to the same speakers who are responsible for the interaction in the AI data; the RMS data of these speakers follow the same pattern as their AI ones. Thus, in /ri/, the statistically significant F-ratio presented in Table 7 is due to HP and KAP (HP: $F(1,8)=25.33$, $p<0.001$; KAP: $F(1,8)=5.94$, $p<0.039$); in /ra/, it is due to KAP ($F(1,8)=8.006$, $p<0.02$). The other subjects do not differentiate between the one- and two-word versions of either /ri/ or /ra/. The complete agreement between AI and RMS

Table 7:

F-ratios and probability levels from 2-way ANOVAs for comparison of AI (right) and RMS (left) in one-word and two-word test phrases.

A * indicates speaker and word makeup interaction.

	<u>AI</u>		<u>RMS</u>	
	<u>F(1,8)</u>	<u>p</u>	<u>F(1,8)</u>	<u>p</u>
a	0.31	n.s.	2.89	n.s.
ri	2.62*	n.s.	6.03*	0.038
sta	1.50	n.s.	2.80	n.s.
<hr/>				
ra	5.20*	0.049	6.44*	0.033
ði	0.23	n.s.	0.49	n.s.
ko	0.54	n.s.	1.10	n.s.

data is not surprising: as there are virtually no durational differences between the vowels⁴ of the one- and two-word versions of the test phrases, the rare differences observed between the AI data of the two versions could only be related to differences in RMS.

In summary, the two members of each test phrase do not on the whole differ substantially in amplitude. The few observed discrepancies hardly alter the general picture, since they concern either one speaker at a time or two speakers who follow opposing trends.

3.3.4 FUNDAMENTAL FREQUENCY

Typical F0 contours for both test phrase pairs are presented in Figures 3 to 6. There are no important differences between the contours of the one-word and the two-word versions of the test phrases. Although there were some minor individual variations among speakers, each one of them followed the same pattern for both members of a pair.

In the /'ari 'stasu/ and /.ari'sta su/ contours, /a/ is rising, /ri/ is high and falling and /sta/ is low and rising in the contours of both one- and two-word versions. In both /psa.raði'ko tus/ and /psa'ra ði'ko tus/, /ra/ is rising, while /ko/ is falling, a natural consequence of the fact that the test phrases are sentence final. The F0 pattern of /ði/ is more variable, but the same within each speaker's data: in those of AA and AP the F0 of /ði/ is either flat or falling, while in the data of

⁴ When relating the durational data to the RMS and AI ones, vowel durations rather than syllable durations are considered, since the RMS and AI measurements refer to the RMS and AI of the syllable nucleus, i.e. the vowel.

HP it is rising and remains high (there are no F0 traces of /ði/ in KAP's tokens). In short, as with duration and amplitude, there are no important differences in F0 contours between the one- and two-word versions of the test pairs.

Figure 1

Means of normalised AI of /arista/. For each speaker the mean of the two-word version is on the left, and the mean of the one-word version on the right.

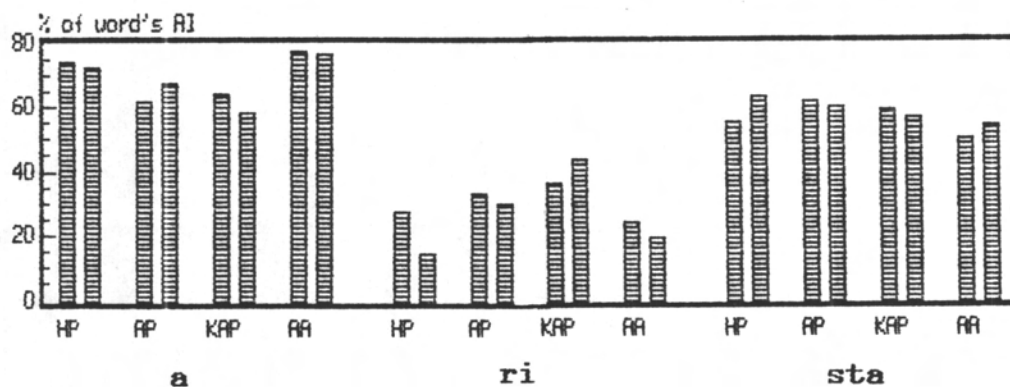


Figure 2

Means of normalised AI of /raðiko/. For each speaker the mean of the two-word version is on the left, and the mean of the one-word version on the right.

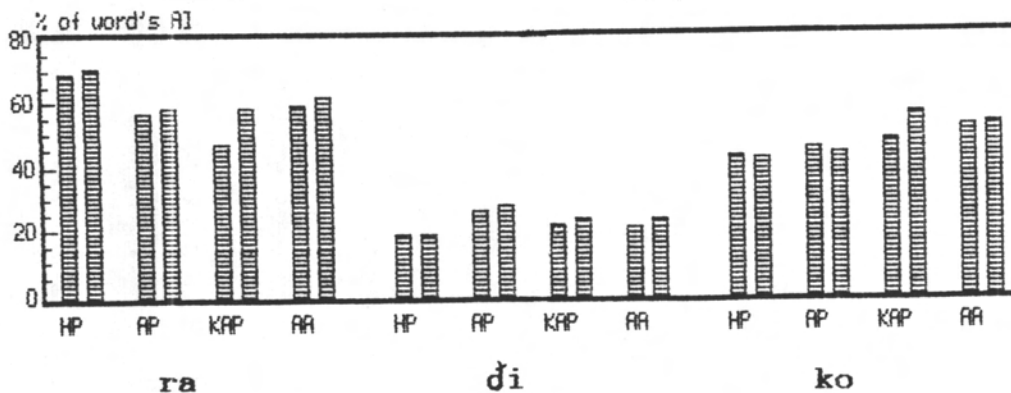


Figure 3

Speaker KAP: Waveforms and smoothed F0 contours of two-word (top graph) and one-word (lower graph) /arista/.

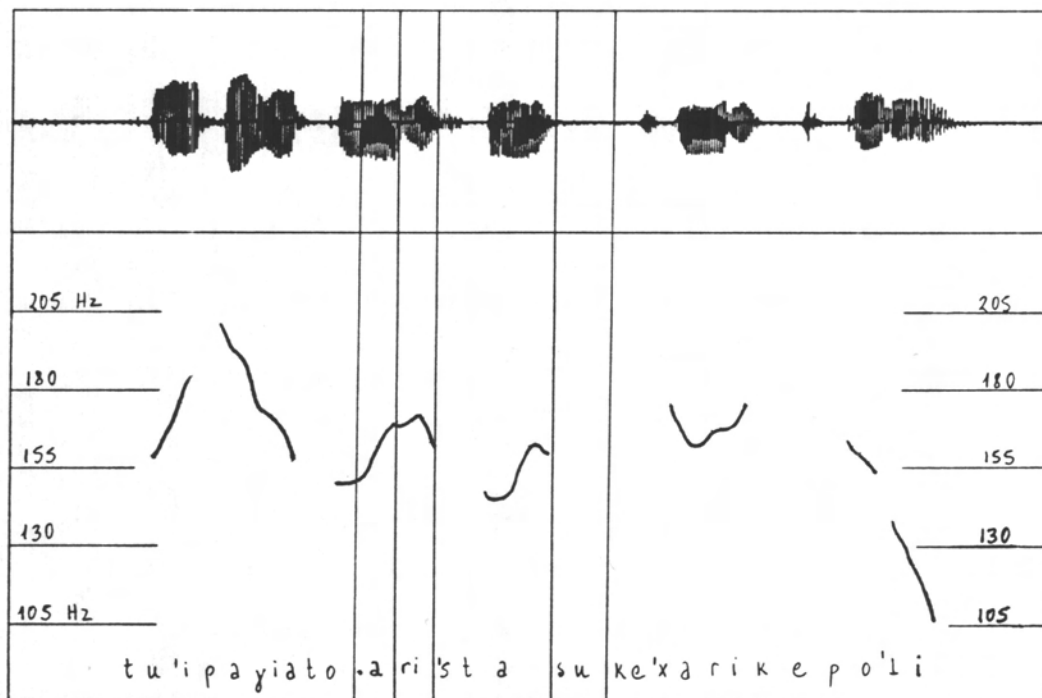
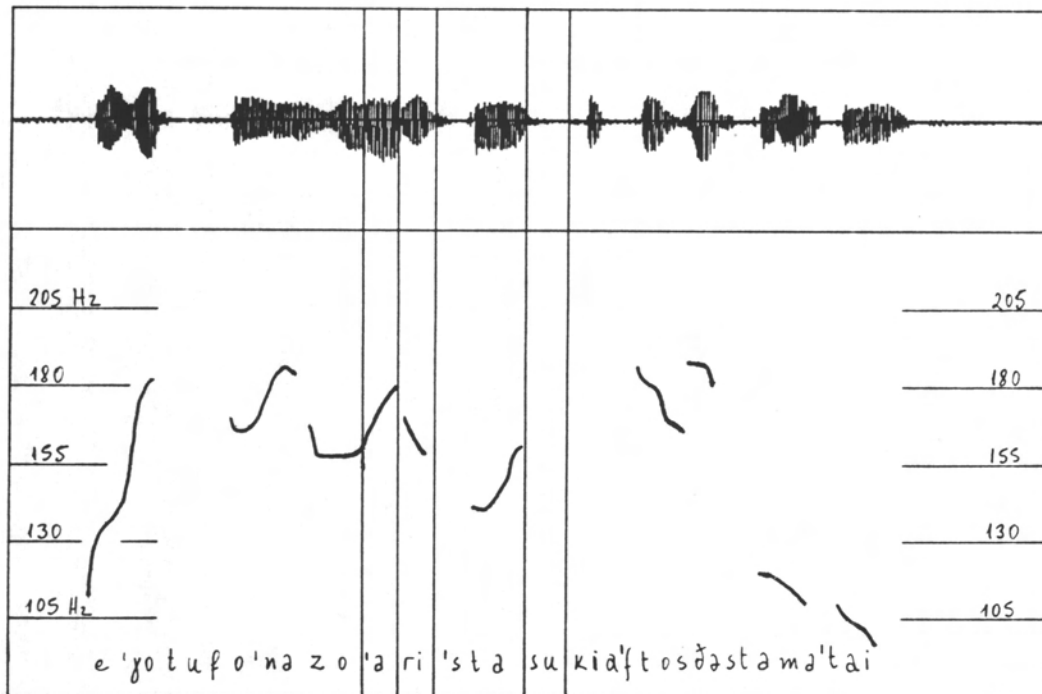


Figure 4

Speaker AP: Waveforms and smoothed F0 contours of two-word (top graph) and one-word (lower graph) /arista/.

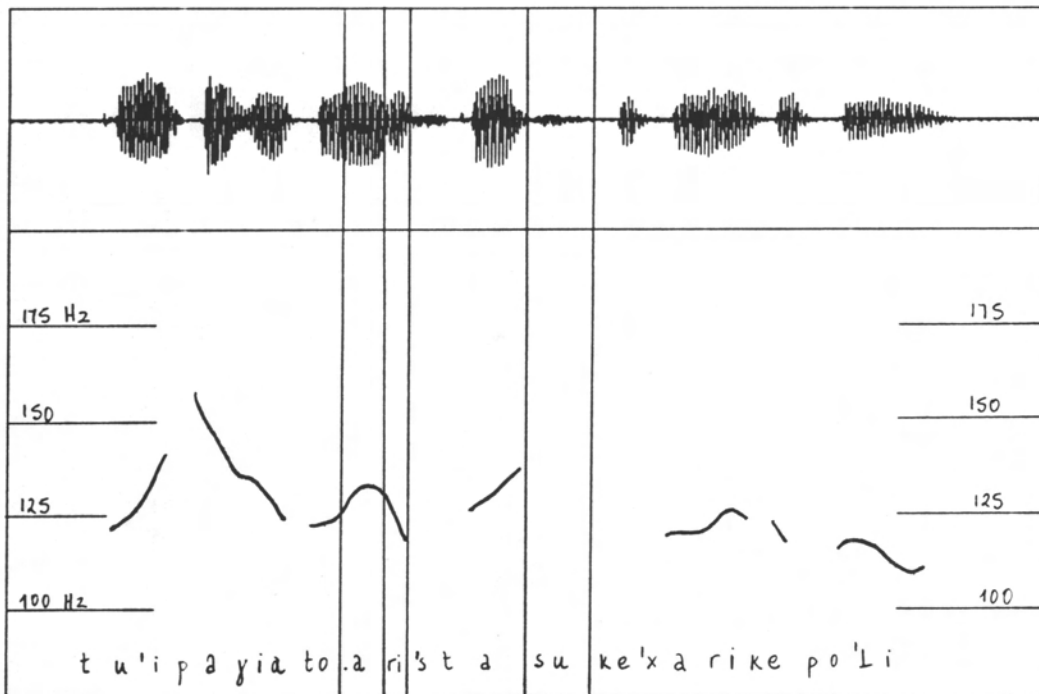
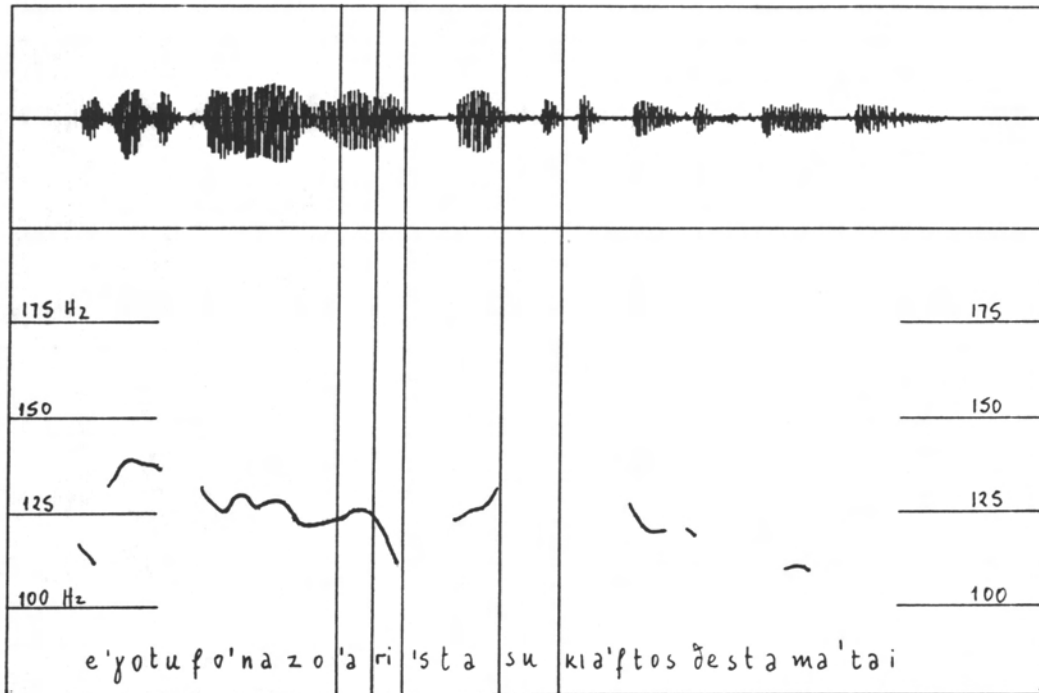


Figure 5

Speaker HP: Waveforms and smoothed F0 contours of two-word (top graph) and one-word (lower graph) /psaraðiko/.

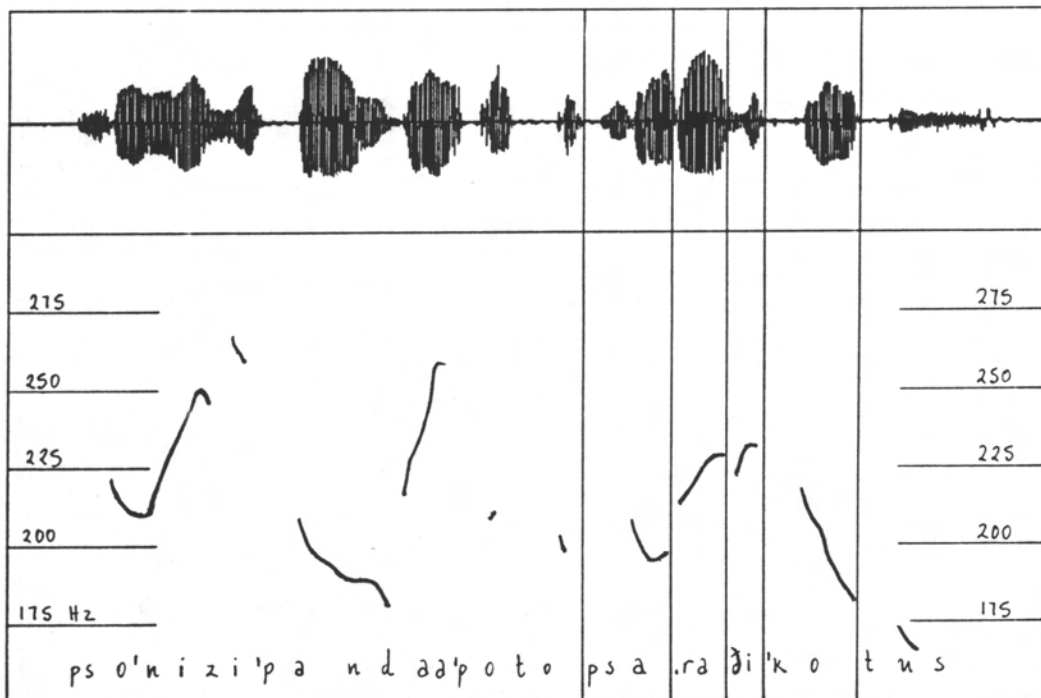
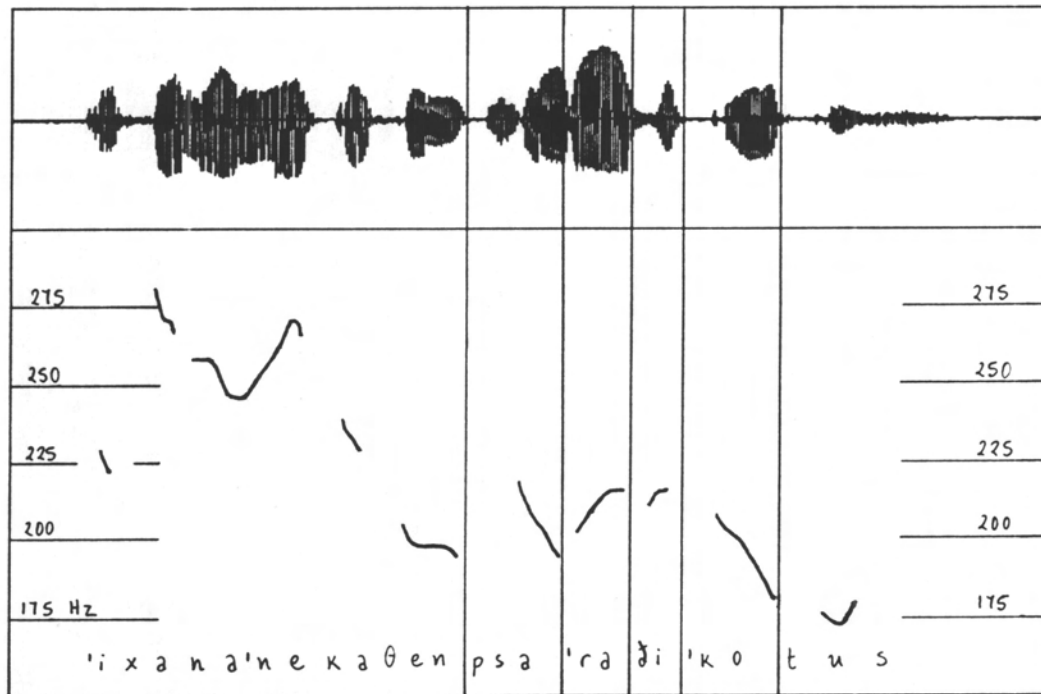
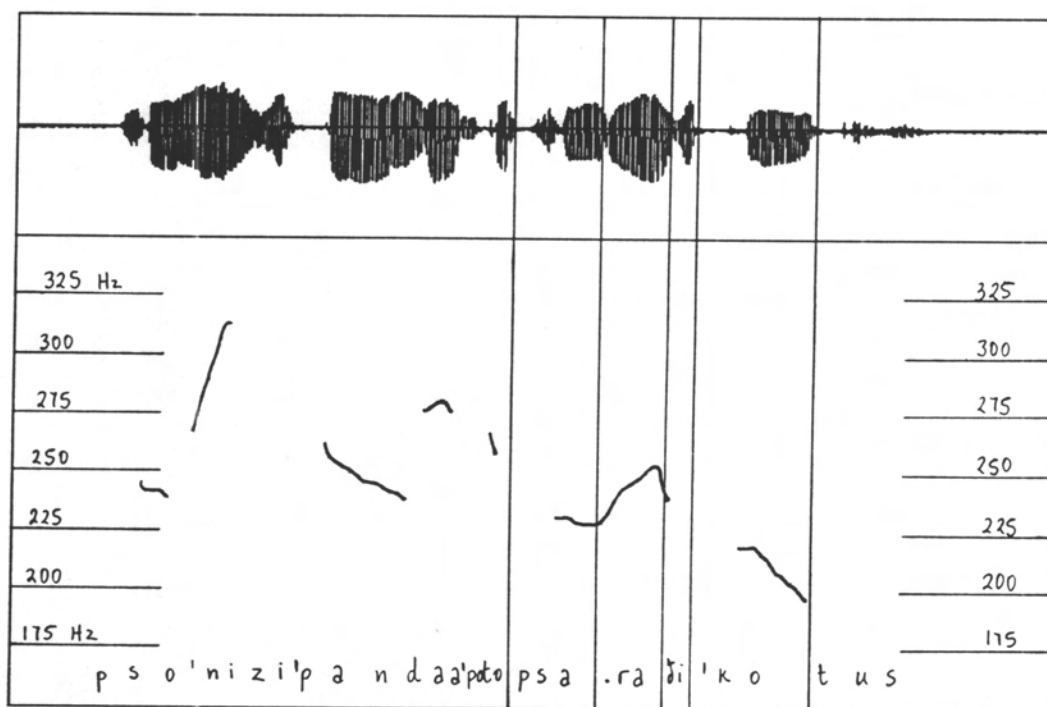
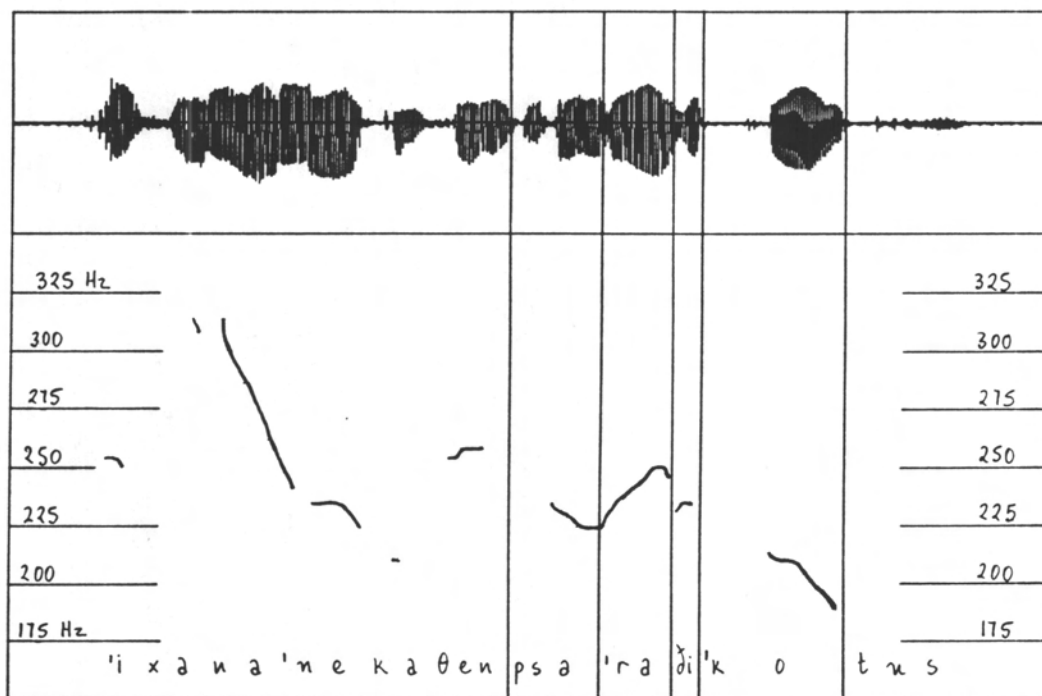


Figure 6

Speaker AA: Waveforms and smoothed F0 contours of two-word (top graph) and one-word (lower graph) /psaraðiko/.



3.4 DISCUSSION

The results of the perceptual test show that native speakers of Greek cannot differentiate between the rightmost lexical stress of a phonological phrase (ϕ) and a SWFC-induced stress when they fall on the same syllable of phonemically identical phrases. The acoustic analysis of the test material confirms that there is no acoustic basis on which such a perceptual distinction could have been made.

On the whole, there are no major acoustic differences between the members of each test phrase. The F0 contours show a remarkable similarity between the one- and two-word versions. The durational and amplitude results, however, contain certain discrepancies. An explanation of the differences concerning the /ri/ syllable of the /arista/ pair could be that the speakers who differentiated between the one- and two-word versions of /arista/ were perhaps somewhat confused as to how to pronounce /'ari 'stasu/ without inserting, metaphorically speaking, quotes to it, and without pausing between the two words as a call would require: "Ari! Stop!". This effect is most evident in HP's data in which the /ri/ of the one-word version is longer than that of the two-word version, while it has lower AI and RMS. HP's data imply that she was probably trying to make a distinction between /arista/ with a word boundary (two-word version) and /arista/ without a word boundary (one-word version) because of the "quotes effect". The fact that the /raðiko/ data are more uniform than the /arista/ ones, corroborates this explanation concerning the discrepancies observed in the /arista/ data.

In /raðiko/, the only acoustic difference between the one- and two-word versions is in KAP's data which show higher AI and RMS on the one-word /ra/. This result indirectly supports Setatos's (1974) and NV's (1986) suggestion that in a host-and-clitic group the stress of the host is stronger than the enclitic stress, by showing that the lexical stress of the host in /psa.raði'ko tus/ is stronger than the subordinate lexical stress of /psa'ra/ in /psa'ra ði'ko tus/. However, this is indirect and weak evidence, since it concerns a single case out of eight (2 test phrase pairs x 4 subjects). Moreover, the fact remains that the one- and two-word versions of both test pairs were perceptually indistinguishable; KAP's stimuli were not easier to distinguish than the rest.

Therefore, the present data do not support Setatos's and NV's suggestion that the lexical stress of the host remains the most prominent in the host-and-clitic group, whereas the SWFC-induced stress is secondary. The SWFC-induced stress may be assumed to be the most prominent stress in a host-and-clitic group, in the same way that the most prominent stress in a ϕ is the rightmost lexical stress (see also section 3.1.1 and Botinis (1989:63 ff.) on acoustic results on this matter⁵). This conclusion agrees with the description of the phenomenon by most studies of Greek, both phonological (e.g. Joseph & Warburton 1987, M-DD) and phonetic (e.g. Dauer 1980a).

The results also indicate that Botinis's proposal that the SWFC-induced stress belongs to a perceptually distinct prosodic category is incorrect. As mentioned in Chapter 1, section 1.6,

⁵ Botinis (1989) shows that in /'maθi'ma tu/ *his lesson* /ma/ with enclitic stress is longer in duration and higher in amplitude and F0 than /ma/ with lexical stress.

Botinis (1989) proposes that enclitic stress forms a separate prosodic category which he calls phrase stress. Botinis's arguments for proposing this category are, however, problematic. His first argument is that phrase stress is syntactically determined (not lexically like word stress) and its function is to delimit phrase boundaries and clarify ambiguous attachment of clitics. But if his proposal were correct phrase stress should appear in every phrase with an enclitic and not only in those which violate SWFC (*Trisyllabic Constraint* in Botinis's terms).

Botinis's second argument is that phrase stress is acoustically distinct from word stress. This argument, however, is very weak: even in Botinis's own data, phrase stress is acoustically so similar to word stress as to prompt Botinis (1989:85) himself to note that "[o]n acoustic grounds it is questionable if there is enough evidence to differentiate between word and phrase stress [...]". The data of Experiment 3 confirm that the two are not acoustically distinct.

Botinis's third argument in favour of phrase stress is that it is perceptually different from word stress. Botinis did two perceptual tests on phrase stress, using synthesis by analysis to create stimuli which were derived from the phrases

(1a) /to .ono'ma mu 'itane γno'sto/ *lit. the name mine was known > my name was known*,

(1b) /to 'onoma mu 'itane γno'sto/ *lit. the name to me was familiar > the name was familiar to me*.

In these tests, stress cues (F0, duration and amplitude) were manipulated one at a time and finally all together to shift the listeners' responses from a structure with one word stress and one phrase stress, as in (1a), to a structure with one word stress only, as in (1b), or vice versa. Briefly, the listeners' responses changed from one to the other of the above phrases only when all three stress parameters in the sequence /onoma/ changed to those of the other form. Changing F0 only, while amplitude and duration remained unaltered, resulted only in a small change of the listeners' responses from (1a) to (1b) (by 41%), or (1b) to (1a) (by 37%). In contrast, in experiments on word stress perception, which were prepared using the same technique, a F0 contour change was sufficient to shift stress perception from (2a) to (2b) or from (2b) to (2a) by 100%.

(2a) /'nomo/ *law*

(2b) /no'mo/ *county*.

Botinis interprets these results as indicating a perceptual difference between word and phrase stress, but in my opinion there is a flaw in his argument. Specifically, the experiments on word stress test whether stress can be associated with one or another syllable within a word, when one stress parameter, F0, is manipulated⁶; i.e. in this case the listeners must compare syntagmatically the two syllables of each word and determine which one sounds more prominent on the basis of F0, while knowing, as native speakers of Greek, that one syllable must be stressed. In the phrase stress experiments, however, the aim is to create the impression of stress on a syllable that was originally unstressed, or to make an originally stressed syllable sound unstressed by manipulating

⁶ As mentioned, Botinis also manipulated duration, average amplitude and amplitude integral in his perceptual tests, but uses only F0 for his argument because F0 is by far the most robust perceptual cue in his tests.

F0. In other words, in this case, the listeners are asked to make a paradigmatic comparison between the syllable they hear and what they know as speakers of Greek to be a stressed (or unstressed, as the case may be) syllable in their language. Moreover, if we assume that the listeners make a syntagmatic comparison among the syllables of /onoma/ in the phrase stress test, another problem arises. In the phrase stress test, the stimuli contain one syllable, the initial /o/, which is stressed in both /'onoma mu/ and /ono'ma mu/: in both phrases /o/ is accompanied by a considerable F0 rise. The stress on /o/ fulfils the requirement of Greek grammar that there is one stress in every word. Therefore, the listeners do not need to look for a second stressed syllable within the /onoma/ sequence. (This suggestion could also explain why a smaller percentage of the listeners responses shifts from (1b) to (1a) than vice versa.) Obviously, the task in the phrase test is much more subtle and difficult than the task in the word stress test, and it is very likely that the listeners found it simply too confusing.

In my opinion, Botinis's findings, and enclitic stress in general, can be accounted for more naturally by a cyclic application of stress in Modern Greek, one cycle taking place lexically and the second one post-lexically. I present here an outline of the implementation of this proposal which is further discussed in Chapter 6, sections 6.5.3 and 6.5.4. Word stress placement is a lexical process, while the enclitic stress in host-and-clitic groups is the result of post-lexical application of the SWFC. This difference becomes clear if one considers again cliticisation and suffixation. Although syllable addition is common to both of these processes, they yield different results: whereas suffixation results in a shift of the main stress as in

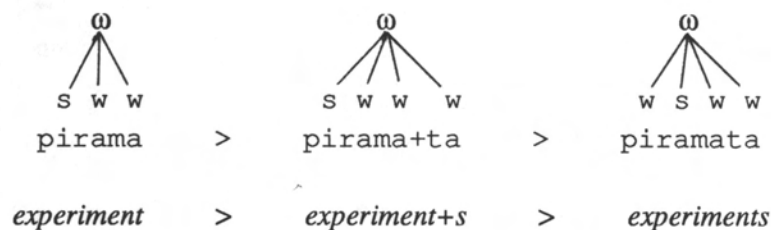
(3) /'maθima/ *lesson* > /ma'θimata/ *lessons*,

cliticisation results in a stress addition, as in

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

This is precisely because suffixation takes place within the lexical component, where the position of stress can be altered, whereas cliticisation is a post-lexical process. On leaving the lexical component, all words, except clitics, form independent phonological words (ω), like the final form in (5).

(5)



The fact that all words constitute independent ωs is true even of monosyllabic "content" words. The difference between these and clitics becomes apparent when one considers examples

like

(6) /'anapso to 'fos/ *turn on the light*

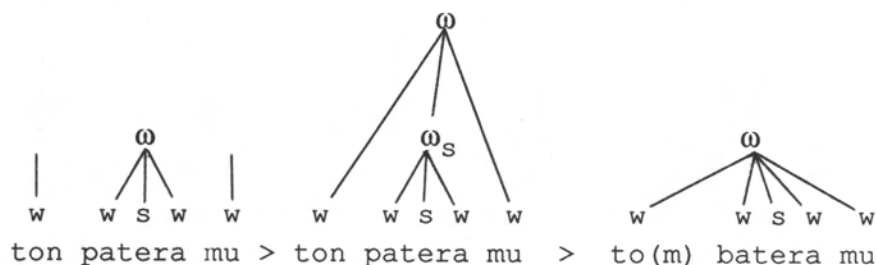
which show that SWFC violations do not arise in sequences of independent words because these form separate ω s.

Clitics, however, remain unattached weak syllables until they are attached to a host post-lexically. Clitics extend the boundaries of ω s by forming compound ω s with their hosts, as in (7b). These compound ω s may be simplified, as in (7c), after sandhi rules related to the clitics, like Nasal Assimilation and Stop Voicing exemplified in (7), have operated.

(7a)

(b)

(c)

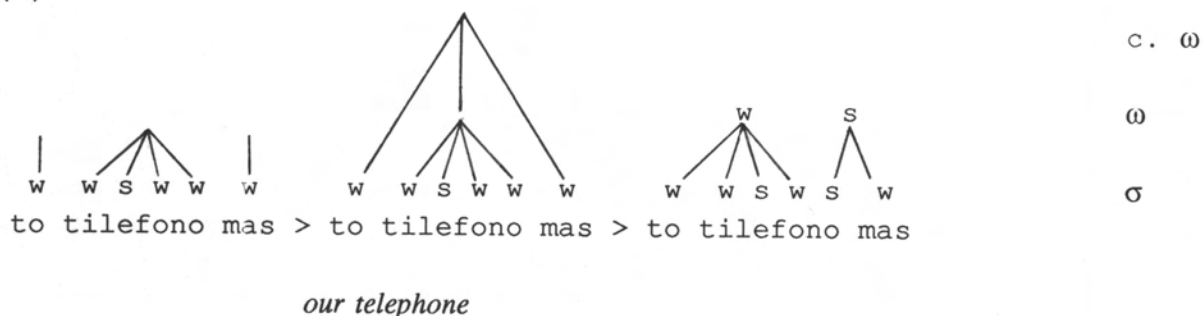


my father (acc.)

The compound ω formed by the host and its clitics still has to conform to the SWFC. When cliticisation does not result in a SWFC violation no change of stress pattern is necessary. When, however, the SWFC is violated by the addition of enclitics, the results of the violation are different from those observed within the lexical component. This is precisely because the host has reached the prosodic component of the grammar as an independent ω , with its own lexical stress. Thus, in SWFC violations, the host's stress cannot move from its position, as it does within the lexical component. The only alternative is for another stress to be added in such a position that it can comply with the SWFC. Hence the stress two syllables to the right of the host's lexical stress. In this case the compound ω cannot be simplified, but is divided in two ω s as in (8)⁷.

⁷ In (8), a slightly different convention in the drawing of the trees has been adopted for reasons of typographical simplicity. This style of representation will be followed in all subsequent trees.

(8)



The advantages of this structure are several. First, it captures the subordination of the host's stress, while at the same time it shows that the two stresses in a host-and-clitic group belong to one ω , albeit a compound one, not to different prosodic categories as Botinis suggests. A second advantage is that this structure can adequately explain sandhi phenomena like those shown in (7).

A third advantage of this structure is that it can easily account for Botinis's finding that in a host-and-clitic group the strongest stress is always the enclitic one. Botinis found that in a phrase like

(9) /to .pro'gra'ma tis/ *her programme*,

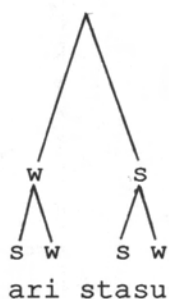
sentence stress (roughly the intonation nucleus) always falls on the phrase stress of the structure, (/ma/), whether both or only one of the host and enclitic are in focus. Leaving aside considerations of whether Botinis is right to assume that either the host or the clitic alone can be in focus, the structure presented in (8) explains why Botinis's expectation is not realised. This is because the two stresses of the host-and-clitic group belong to *one* ω . Within this ω , it is not possible for the host's stress to become more prominent than the enclitic one for two reasons. First, because such a change of prominence would violate the SWFC; second, because the relationship of the two stresses in a host-and-clitic group is analogous to that of the two stresses in an English word like *familiárity*. Expecting that the secondary stress of a word like *familiárity* could become more prominent than the word's primary stress not because of rhythmic requirements, but because the focus has shifted on a different element within the word is clearly nonsensical. The same situation holds for the Greek compound ω s.

The present proposal, however, has two possible disadvantages. First, in structures like (8) there is no motivation for choosing between /fonomas/ and /nomas/ as the second constituent of the compound ω . For instance, there are no phonological processes, such as sandhi, taking place between /fo/ and either the preceding /le/ or the following /no/ (or the equivalent syllables of other words) which would justify linking /fo/ with either ω ; similarly there are no phonological processes which do *not* take place between /fo/ and either /le/ or /no/, an absence which would also allow /fo/ to be a constituent of only one or the other ω .

It was the lack of sandhi rules operating in this environment that prompted the study of the middle syllables of /arista/ and /raðiko/. Given that the branching in the two-word versions is

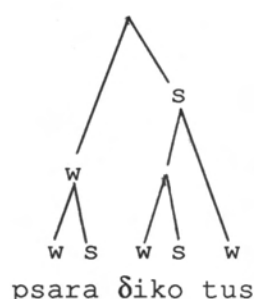
known, since it is determined by word boundaries, possible acoustic differences between the /ri/ and /ði/ of the one- and two-word versions could indicate a difference in branching.

(10a)



Ari stop

(b)



their own fishmonger

ϕ

c. ω

ω

σ

Specifically, as can be seen in (10a) and (10b), in the two-word /'ari 'stasu/, /ri/ belongs to the first ω , while in the two-word /psa'ra ði'ko tus/, /ði/ belongs to the second ω . Thus, if in the one-word versions of the two test phrases the middle syllable belongs to the first ω , the one- and two-word version /ri/s should be acoustically similar, while the one- and two-word version /ði/s should be acoustically different. If, on the other hand, in the one-word versions the middle syllable belongs to the second ω , the reverse trend should be observed. The data show some differences between the one- and two-word /ri/s and no differences between the one- and two-word /ði/s, suggesting that the syllable between the two stresses is a constituent of the second ω .

On the other hand, the differences between the one-word and two-word /ri/s were observed only in the data of two of the speakers and, as mentioned, they may be due to the fact that these speakers could not read the two versions in exactly the same way, due to their inherent differences (/ 'ari 'stasu/ could have been interpreted as a call with a pause between the two words). If this is so, then the data indicate that either branching of the syllable between the two stresses is acceptable. This matter will be discussed further in Chapter 6, section 6.5.4.

The second apparent disadvantage of the present proposal is that the ω s which compose a compound ω do not correspond to any particular morphological or syntactic constituent. This sets them apart from the standard ω of Greek which, according to NV (1986), corresponds to the terminal element of the syntactic tree. In my opinion, this is a superficial problem, since the ω s which are formed according to NV's (1986) rules are the product of the lexical component, whereas the constituent ω s of a compound ω are formed post-lexically. In other words, the ω s which reach the prosodic component of phonology from the lexical component are formed according to the criteria set by NV (1986), with the exception of clitics which, as mentioned, leave the lexical component as weak syllables; post-lexically, when the clitics are attached to their hosts, the boundaries of ω s are allowed to change to accommodate the requirements of sandhi rules, prominence patterns and so on. (The questions of constituent structure and domain

boundaries in Greek will be discussed further in Chapter 6, section 6.5.4.)

3.5 CONCLUSION

Experiment 3 aimed at examining the perceptual and acoustic nature of enclitic stress. The data unequivocally show that the enclitic stress is the most prominent stress in a host-and-clitic group. Enclitic stress is perceptually identical to a lexical stress, while the weakened lexical stress of the host is acoustically and perceptually similar to a subordinate lexical stress. To account for enclitic stress, a solution is proposed according to which the Stress Well Formedness Condition applies both lexically, moving lexical stress to the right of its original position, and post-lexically, adding a stress two syllables to the right of the host's stress in a host-and-clitic group which violates the SWFC. The two stresses in a host-and-clitic group are represented by means of a compound ω .