

**Second language acquisition of intonation:  
the case of Dutch near-native speakers of  
Greek**

**Ineke Mennen**

Thesis submitted for the degree of Doctor of Philosophy  
University of Edinburgh

July 1999



I hereby declare that I have composed this thesis,  
and that the work in it is my own unless explicitly  
stated otherwise.

Signature:

## Acknowledgements

First of all, I would like to thank my supervisors Bob Ladd and Antonella Sorace for setting aside some of their precious time. In particular I am grateful to Bob Ladd, for his enthusiasm, advice, and for giving me the opportunity to come to Edinburgh. Without this offer, I would never have started a PhD.

Secondly, I am grateful to many people in the Department of Linguistics. Ethel Jack, who was always there with a cup of coffee and a chat, and who always managed to cheer me up when things seemed difficult. Irene MacLeod, who kept me up to date about the latest developments in the department. Eddie Dubourg, who provided unrelenting computer support and patience. Cedric MacMartin for answering all my questions about xwaves.

A huge vote of thanks goes to the 6 Dutch speakers of Greek, without whom this project would not have been possible. It requires a lot of confidence and courage to let someone record and analyse your speech, especially in a foreign language. Don't get discouraged by the mistakes I found... your Greek is excellent! Thanks are also due to the Greek and Dutch native control speakers for their time and assistance, especially to Dora and Antonis who participated in all the experiments.

Of course, I also want to thank my fellow students and colleagues. Max, for putting up with my moaning and insecurities; Astrid, for helping me out with statistics, and most of all for being so considerate and fun to be with; Dora, for always being there and showing me the way; Heather, for being a friend; Alice, for reorganising my thoughts and for running around the Meadows. Thanks also to my office mates Catriona, Etsuko, Satsuki, Patricia and Lynda. I am also grateful to my colleagues at QMUC, it was such a nice change from working on my PhD.

Thanks also to my parents, who unquestioningly accepted my sudden change of career, and supported me throughout.

Finally, I would like to thank my friends in Holland. I especially want to thank Ron, who told me I would always regret it if I didn't do a PhD; Paula, for giving me a different perspective to life and making that lovely mistake in 'aluminium pannen'; and Lysbeth, for giving up her room every time we visited Haarlem. I am especially grateful for the great time we all spent together at that lovely beach in Greece.

The most grateful I am to Petto, who was always there to support me and has had to endure my mood swings during the whole PhD process. You always believed I could do it, and here I am!

# Abstract

This thesis focuses on the second language (L2) acquisition of intonation by Dutch near-native speakers of (Modern) Greek. Five experiments were carried out which can be divided into two main areas: one which was concerned with peak alignment (i.e. the timing of the intonational peaks with respect to the syllables), and another which focused on nucleus placement (the placement of the most prominent word in the utterance).

The first three experiments were concerned with peak alignment in Greek prenuclear accents. A first experiment showed that there are cross-linguistic differences in the alignment of the peak in Greek and Dutch. In Greek prenuclear accents the peak is aligned just after the onset of the first postaccentual vowel. In Dutch, however, the peak is aligned earlier, and is influenced by the phonological length of the accented vowel of the word bearing the prenuclear accent. (Specifically, the peak is at the end of the accented vowel when that vowel is long, but at the end of the following consonant when it is short.)

The next experiment showed that most non-native (Dutch) speakers of Greek failed to exhibit native-like peak alignment in the L2. The speakers aligned the peak in their Greek data as early as they would in Dutch when the test word has a long vowel in the accented syllable. It was also shown that speakers did not develop a 'merged' system, with values half-way between the monolingual norms for Dutch and Greek.

Experiment 3 examined the effect of L2 learning on the peak alignment in prenuclear accents in the first language (L1). It was shown that even though most speakers did not master the alignment values in the L2, the alignment pattern in their L1 was nevertheless affected.

The last two experiments were concerned with another aspect of intonation, namely the realisation of nucleus placement in Greek yes/no questions. The yes/no questions were of two types: (i) nucleus-final (NF), in which the nucleus or main stress was located on the utterance-final

word, and (ii) nucleus-non-final (NNF), in which the main stress was on an earlier word in the utterance. It was shown that the choice of nucleus placement was problematic for the group of non-native speakers. Only half of the non-native speakers produced the two different nucleus locations. Furthermore, non-native speakers' production of both types of yes/no questions differed from that of native speakers of Greek, specifically with respect to scaling (i.e. the fundamental frequency at which a certain pitch peak or valley occurs). However, their production of NF yes/no questions was closer to the native Greek norms than was their production of NNF yes/no questions.

An attempt was made to interpret the results of the experiments in the light of Flege's Speech Learning Model (SLM), which was originally developed to account for segmental aspects of L2 learning. It was found that the ability of the model to account for intonational data was limited, and more specifically, that the notion of 'phonetic similarity' needs to be further refined.

**To Ro daSilva**

## Table of contents

<b>Chapter 1: Introduction</b>	<b>1</b>
1. General introduction	1
2. Development of second language acquisition research	3
2.1. Contrastive analysis	3
2.2. Changing views on transfer	5
2.3. Interlanguage phonology	9
3. Factors that affect speech production and perception in L2	14
3.1. Native language influence	14
3.2. Types of phonetic categories and contrasts	16
3.3. Age of learning	17
3.4. Language experience and ultimate attainment	18
3.5. Effect of training	21
4. Current second language speech models	22
4.1. Criteria for measuring phonetic similarity	27
5. Second language acquisition of intonation	28
5.1. An overview	28
5.2. Describing intonation across languages	32
5.3. Possible errors in L2 intonation	34
5.3.1. Alignment	34
5.3.2. Pitch range and tonal scaling	38
5.3.3. Word stress and nuclear accent placement	41
6. Aims and overview of the thesis	43
<b>Chapter 2: Intonation in Dutch and Greek</b>	<b>46</b>
1. Dutch intonation	46
1.1. The Grammar of Dutch Intonation (GDI)	51
1.1.1. Criteria for classification	51
1.1.1.1. Rate of change	52
1.1.1.2. Size of pitch movement	54

1.1.1.3. Timing	55
1.1.2. Pitch contours	56
1.2. Alignment in Dutch	56
1.3. Shortcomings of alignment studies	61
1.4. Prenuclear accents in Dutch	62
1.5. Yes/no question intonation in Dutch	64
2. Greek intonation	68
2.1. Prenuclear accents in Greek	71
2.2. Greek yes/no questions	74
3. Comparison of Dutch and Greek intonation	77
3.1. An autosegmental-metrical analysis of Dutch and Greek intonation	77
3.2. Flege's notions of 'new' and 'similar'	85
<b>Chapter 3: Prenuclear accents</b>	<b>90</b>
1. Introduction	90
2. Experiment 1	90
2.1. Introduction	92
2.2. Method	94
2.3. Results	102
2.4. Summary of results	108
3. Experiment 2	111
3.1. Introduction	111
3.2. Method	113
3.3. Results	118
3.3.1. Group analyses	119
3.3.2. Individual peak alignment analyses	121
3.3.3. Additional group analyses	123
3.4. Summary of results and discussion	125
4. Experiment 3	127
4.1. Introduction	127
4.2. Method	128
4.3. Results	130

4.4. Summary of results and discussion	136
5. Rate and duration	138
6. General discussion	141
<b>Chapter 4: Nucleus placement in L2 Greek yes/no questions</b>	<b>146</b>
1. Introduction	146
2. Experiment 4: nucleus non-final yes/no questions	148
2.1. Introduction	148
2.2. Method	149
2.3. Methodology	155
2.4. Results	159
2.4.1. Group results	160
2.4.2. Individual speakers' results	166
3. Experiment 5: nucleus final polar questions	172
3.1. Method	173
3.2. Results	175
3.2.1. Group analyses	176
3.2.2. Individual analyses for speaker DG3 and DG5	181
3.3. Summary of results	182
4. Discussion	186
<b>Chapter 5: General discussion</b>	<b>189</b>
1. Overview of the main findings	189
1.1. Alignment	189
1.2. Nucleus placement and scaling	189
2. General discussion	193
2.1. Phonetic similarity and the unit of analysis	193
2.2. Bidirectional influence	197
2.3. Limits on attainment	199
2.4. Types of phonetic categories	200
<b>References</b>	<b>204</b>

Appendix A	222
Appendix B	226
Appendix C	230
Appendix D	235
Appendix E	245

# Chapter 1

## Introduction

### 1. General introduction

Sometimes, when we hear non-native speakers talk in our native language, their speech - apart from obvious mistakes - sounds somehow *different* from that of native speakers. This difference is often referred to as 'accented speech' or 'foreign accent', and we can often perceive it even when the non-native speakers are reasonably proficient in our native language.

It is a common belief that mastering this typical accent of a foreign language is extremely difficult when learning starts in adulthood. Many adult second language (L2)<sup>1</sup> learners will agree with Scovel's conclusion (Scovel, 1969) that it is impossible for adults past puberty to achieve perfect pronunciation in a foreign language.

Researchers have often assumed that accented speech is mainly caused by deviations at the segmental level (i.e. vowels and consonants), and as a consequence most research is limited to this level only. It is, however,

---

<sup>1</sup> In this thesis I will use the term 'second language' or L2, to refer to any language learned later in life, that is after puberty, regardless of where it was learned (in the environment of the first language, or that of the L2), unless explicitly stated otherwise. The reason for not distinguishing between 'second language' (language learned naturally in the environment of the L2) and 'foreign language' (language learned by formal instruction in the L1 environment) is that usually learners experience both kinds of learning to a certain extent, so that it is often not possible to distinguish between the two. Therefore, I will also not make a distinction between 'learning' (which is the process involved when language is learned through formal instruction) and 'acquisition' (which is the process involved when language develops naturally within the L2 environment), a distinction introduced by Krashen (1981).

generally accepted that prosodic factors (like intonation, rhythm, or stress) can contribute to a foreign accent and the interest in the role of these factors in accented speech is growing. In the next sections, previous research in the area of second language phonology will be discussed. A brief overview of the development of this field will be given, followed by a description of factors that affect L2 speech production, and a discussion of two current L2 speech models. As a consequence of the lack of studies dealing with the prosodic level of L2 speech, these sections are mainly based on research at the segmental level. The prosodic aspects of L2 speech will be discussed in section 5 and onwards.

The aim of the discussion in this Chapter is not to provide an exhaustive description of the literature, but rather to present a representative sample of work on L2 phonology and its development. Furthermore, in this overview I will not make an overt distinction between phonological and phonetic levels of analysis, as it is often not easy to distinguish between the two in the course of the development of the field.<sup>2</sup>

---

<sup>2</sup> In many studies in this field no distinction is made between phonology and phonetics, and the term L2 phonology is often used to refer to both phonological and phonetic aspects of second language learning. Many text books on second language acquisition do not explicitly distinguish between the two. Ellis (1994), for example, treats pronunciation and phonology as one subject in the subject index of his book. Sharwood Smith (1994) also does not distinguish between the two and uses the term phonology to refer to both phonetics and phonology. Leather and James (1996) in a recent review note that there may be differences between the two but they state that they "...have felt, however, that it is both possible and preferable in parts of this review to avoid making overt distinctions between phonetic and (inter-) phonological, or between different "phonetic" levels of analysis..." (p.276).

## 2. Development of second language acquisition research

### 2.1. Contrastive analysis

The role of the native language in second language acquisition has been debated for a long time, and there is still no consensus as to the nature of cross-linguistic influences. The roots of the field of second language acquisition can be found in the 'contrastive analysis' approach (Fries, 1945; Lado, 1957), which was in its heyday in the 1950s and 1960s. Until the late 1960s it was generally assumed that native language influences could greatly affect the performance in a second language. Thus, contrastive analysis, where systematic comparisons of the native language and the second language were made, was seen as very useful to explain problems in language learning, and to develop teaching curricula. The proponents of the contrastive analysis approach (e.g. Fries, 1945; Lado, 1957) claimed that there was a strong relation between the amount of similarity of the first language (L1) and the L2, and the degree of difficulty of learning. The influence of the native language on the learner's production and perception of a second language was seen within a behaviorist framework. Within this framework, first language acquisition was seen as the learning of a fixed set of habits. In second language acquisition, it was assumed that the learner would transfer the linguistic habits of the L1 to the L2. This process of 'language transfer' could be either positive or negative. Where first and second language are similar, positive transfer might take place; but where they are different, the habits of the L1 will interfere with those of the L2 (negative transfer).

In the early 70s, the contrastive analysis hypothesis began to fall into disrepute. There were a number of reasons for this reaction to the contrastive analysis hypothesis. First of all, it was a reaction against the claims of Fries (1945) and Lado (1957) about the predictive power of contrastive analysis. Lado (1957) claims that it could predict when interference errors would occur in second language acquisition.

....the student who comes in contact with a foreign language will find some features of it quite easy and others extremely difficult. Those elements that are similar to his native language will be simple for him, and those elements that are different will be difficult. The teacher who has made a comparison of the foreign language with the native language of the students will know better what the real learning problems are and can better provide for teaching them (Lado, 1957: p.2).

However, empirical evidence proved otherwise: many cases of positive transfer failed to occur; learners failed to exhibit the errors predicted by negative transfer; and some errors appeared which could not be attributed to the native language. A second reason for the rejection of the contrastive analysis approach was its association with behaviorism. Behaviorist and structuralist views, and as a consequence contrastive analysis, became very unpopular in the late 1960s, when ideas about the innateness of language started to gain ground (Odlin, 1989).

As a result, researchers turned their attention to 'error analysis', which analysed and classified the errors of second language learners. The error analyses provided a wide range of evidence for errors which could not be explained by language transfer. Some errors seemed to arise from other sources, like overgeneralisations (inappropriate use of a second language rule); and simplifications (e.g. the omission of articles). Furthermore, many errors could be attributed to developmental factors. In this period, empirical research showed a similarity of errors in the acquisition process among language learners of different backgrounds, and among both first and second language acquisition (Dulay and Burt, 1974; Flege, 1980; Hecht and Mulford, 1987). As a consequence, the credibility of contrastive analysis was damaged even further. However, error analysis also failed to provide convincing theoretical conclusions about the mechanisms of second language acquisition. For example, it could not sufficiently distinguish transfer errors from errors due to developmental processes.

At the time when the contrastive analysis hypothesis and the field of error analysis had lost most of their theoretical ground, the attitude towards language acquisition in general was shifting as well. Scholars started to become more and more interested in the notion of language

universals and a predisposition for language learning. The difficulty that most adults experience in learning a second language was thought to be biologically determined. The idea of biological factors governing language acquisition emerged from a study of Penfield and Roberts (1959). They attributed the age-limitation for language acquisition to the loss of plasticity of the brain. Lenneberg (1967) refined this critical period hypothesis further by linking it to a neurological process, called cerebral lateralisation for language functions. According to Lenneberg, it was impossible to fully acquire a language after lateralisation is completed, i.e. after puberty. This explanation for the existence of a critical period for language learning has since been largely disproved. Krashen (1973), for example, presented evidence that cerebral lateralisation for language processing may be complete long before puberty, in fact, by the age of five or even earlier. As a result, most scholars started looking for explanations other than lateralisation, and the original critical age hypothesis was revised.

## 2.2. Changing views on transfer

As mentioned in the previous section, transfer was initially seen within a behaviorist framework of learning. It was assumed that learners of a second language would carry over the 'habits' of their native language to the L2. The degree of difficulty of learning was considered to depend upon the extent to which the two languages differed from each other. 'Positive' transfer would occur in cases where the native language and the second language showed similar patterns. In cases where the two languages differed, this would result in 'negative' transfer. In other words, positive transfer would facilitate learning of the L2, whereas negative transfer would interfere with it. For example, when a sound such as /p/ or /b/ exists in both the native and the second language (e.g. in Hebrew and English, cf. Selinker, 1992) it is predicted that positive

transfer will occur. In this case the L2 sounds that are identified with sounds in the native language will be substituted by the L1 sound (Selinker, 1992: p. 104). Phonetic interference, or negative transfer, will occur when a sound such as /ð/ does not exist in the native language (a difference between e.g. Dutch and Modern Greek). The Modern Greek /ð/ is likely to be interpreted through the 'phonological filter' of the native language, and will be substituted by the closest L1 sound, e.g. /d/<sup>3</sup>.

The studies on transfer were carried out by means of elaborate contrastive analyses of the native and second languages. Thus, it was believed that the pattern of L1 interference could be predicted. These predictions, however, were not tested empirically until the late 1960s. When the predictions of the contrastive analysis hypothesis were not borne out by the results of subsequent error analysis studies, earlier thinking on transfer was called into question.

As a consequence, views about language transfer have undergone a considerable change. Consider, for example, the 'working definition' of the term transfer, as proposed by Odlin (1989: p. 27)

Transfer is the influence resulting from similarities and differences between the target language and any other language that has been previously (and perhaps imperfectly) acquired.

It is quite obvious that this view on transfer is far removed from its original association with behaviorist theories of L2 learning. Recently, a number of researchers have started to reconsider the importance of transfer in second language acquisition. According to Odlin (1989) the consensus about the importance of language transfer grew throughout the 1980s. He comments:

Despite the counterarguments ... there is a large and growing body of research that indicates that transfer is indeed a very important factor in second language acquisition (Odlin, 1989: p. 4).

---

<sup>3</sup> See also Koutsoudas and Koutsoudas (1983) for segmental errors of Modern Greek learners of English; and Selinker (1992: p.105) for a similar example of segmental errors in Hebrew-English.

Currently, it has been recognised that transfer is a complex process and it is likely to interact with other processes involved in second language acquisition (Ellis, 1994; Hecht and Mulford, 1987; Odlin, 1989). Once it was acknowledged that other factors also frequently play a role in second language acquisition, researchers increasingly started to concentrate on the conditions under which transfer is likely to occur (Odlin, 1989). Some of the conditions that may promote and inhibit transfer are language level, developmental factors, and markedness, which will be discussed below. It should be noted that these are not the only factors, but see Ellis (1994: Chapter 8) for an in depth discussion of other factors, such as social factors, prototypicality, language distance and psychotypology.<sup>4</sup>

Language level refers to the different aspects of language, such as phonology, syntax, semantics and discourse. Although there is evidence for transfer in all aspects of language, it is generally recognized that transfer is more evident at the level of the sound system than at the level of syntax (Beebe, 1987; Broselow, 1988; Scovel, 1969). There is an abundance of evidence on the existence of 'foreign accents', and it is widely accepted that native speakers can easily infer the language background of different learners. However, although the effects of the L1 on the pronunciation are obvious, it is not as simple a process as was previously thought at the time of contrastive analysis. The view that the bigger the difference between the L1 and the L2, the more likely that a foreign accent exists - a view which was held by the contrastive analysis - appears to be inadequate (as will be discussed in section 2.3 and 3.1).

Another process which is likely to interact with transfer is the developmental factor. Developmental factors in L2 language acquisition can be seen from different perspectives. Firstly, it can be seen as the extent

---

<sup>4</sup> Ellis (1994) defines social factors as "the effect of the addressee and of different learning contexts on transfer", prototypicality as "the extent to which a specific meaning of a word is considered 'core' or 'basic' in relation to other meanings of the same word", and language

to which transfer is evident at different levels of development. Corder (1978) introduced the idea of a 'restructuring continuum'. He suggested that the learner's L1 is the starting point of L2 acquisition. As acquisition proceeds, learners gradually restructure their language, and the L1 is gradually replaced by the target language. If Corder's theory is correct, transfer should be more evident in the early than in the later stages of development. Phonology is one of the aspects of language in which transfer seems to be a starting point. For example, Wenk (1986) reported on the acquisition of English rhythm by French learners. He noticed a development in the L2 acquisition of rhythm. Beginners simply transferred the French rhythm into English; advanced learners produced standard English rhythm, while intermediate learners produced a kind of hybrid rhythm system with characteristics of both French and English rhythm. Thus, he observed three developmental stages in the acquisition of L2 phonology, starting from a simple transfer. Major (1986) in his study on the production of L2 English by Brazilians, also found evidence in support of transfer at the starting point of development. However, not all transfer errors appear at early stages of development. Kellerman (1983) observed that certain transfer errors (involving pronominal copies in relative clauses) appeared only at a more advanced stage of L2 acquisition, when the learner had a better knowledge of the target language. Also, transfer errors sometimes fail to be eliminated at later stages of development. Some errors that are clearly traceable to L1 influence have been found to appear in the speech of advanced learners. Flege (1980), for example, found a direct influence of phonetic characteristics of Arabic on English stops produced by advanced Saudi Arabic learners of English. Furthermore, the results of recent studies suggest that even after long exposure to the L2, most L2 speakers never fully master the L2 at the phonetic level (Flege, 1980, 1981, 1987b).

---

distance and psychotypology as "the perceptions that speakers have regarding the similarity and difference between languages" (p. 315).

Another perspective of the developmental factor is the complex interaction of natural principles of L2 acquisition and transfer. For example, Wode (1976; 1980) found that negative transfer seemed to affect some phonological elements, while others were acquired with no influence from the native language, but rather in the same way that a child would acquire them in L1 phonology.

Transfer might also be affected by markedness. In many second language acquisition studies, markedness is defined in the following way: a feature that is present in most languages is considered unmarked, while features that are specific to a particular language are marked. Unmarked features in L1 are more likely to be transferred than marked ones, especially if the corresponding feature in L2 is marked (e.g. Eckman, 1987; Zobl, 1984). Eckman (1987) for example, uses markedness to explain, for example, why voiced obstruents are acquired in word final position after they are mastered in other positions.

### 2.3. Interlanguage phonology

With these new views on transfer and the recognition that other factors than transfer alone may play a role in second language acquisition, the contemporary field of second language research was born. At the time when Lenneberg's critical period hypothesis was called into question, and the contrastive analysis hypothesis was fading, Selinker (1972) introduced the interlanguage theory. Selinker used the term interlanguage to refer to the transitional linguistic systems learners build on their way to full L2 competence. Each intermediate system reflects the learner's current L2 knowledge. It is thought to be separate from both L1 and L2, with its own characteristics and rules. The learner's deviant phonetics is no longer seen, as proponents of the contrastive theory maintained, as the result of simple substitution of some second language sounds with similar L1 sounds. Instead, the interlanguage theory's central idea, that the learner's

grammar is autonomous in nature, was also applied to the study of second language phonetics. In other words, the phonetic deviation from the L1, is not due to interference between the L1 and the L2, but is the output of an interlanguage (Flege, 1980). However, it was not until the mid 1980s that systematic research into interlanguage phonology began. The term interlanguage phonology, or second language phonology, is usually used to refer to both segmental and suprasegmental aspects of L2 phonology and phonetics. Nevertheless, the majority of research in the area of interlanguage phonology has focused on the production and perception of segments (vowels and consonants), whereas the effects upon prosodic features (like rhythm, stress and intonation) has been widely neglected.<sup>5</sup>

The development of interlanguages is not a static process. Interlanguages evolve over time, i.e. they are variable. L2 learners construct a series of interlanguages, and revise them continuously until they have reached the final L2 stage. However, it is recognized that there is a limit to the development of interlanguages. The vast majority of second language learners will never reach the same level of competence as native speakers of that language. Selinker (1972) proposes the term '*fossilization*' for this phenomenon. Fossilizable structures, according to Selinker, are well-known errors like the use of French uvular /r/ [ʁ] in English interlanguage, or American English retroflex /r/ in French interlanguage (Selinker, 1972). Scovel (1969) is convinced that phonological fossilization is inevitable for adult L2 learners. He has even gone as far as offering dinner to anyone who can produce an L2 speaker (who started learning the L2 after puberty) without any trace of foreign accent (Tarone, 1987). Flege (1980) and Flege and Hillenbrand (1984) also found that most experienced L2 learners are only partially successful in

---

<sup>5</sup> In this thesis I will use the term 'prosody' and 'suprasegmental' to refer to "those aspects of speech that involve more than single consonants or vowels" (Ladefoged, 1993: p.243).

producing L2 sounds. Flege and Hillenbrand (1984) describe this phenomenon as a 'merged' system in interlanguage phonology. L2 learners will merge the phonetic properties of the L1 and L2 phones. French and English stops /p,t,k/ are acoustically different sounds, as the French stops have a shorter voice onset time (VOT) than the English stops.<sup>6</sup> So, in the French and English words *tous* and *two*, French learners of English (or English learners of French) will produce /t/ with a VOT value that is intermediate in degree between those observed for the typical French /t/ and the typical English /t/. Flege and Hillenbrand's results suggest that even after long exposure to L2, most L2 speakers will never reach the target VOT value of the L2. Furthermore, the learning of a second language can affect the production of sounds in the L1 as well. Native French speakers, who were experienced speakers of English (12.2 years of residence in an English-speaking environment), were reported to produce French /t/ with VOT values that were intermediate between those of French and English monolinguals. Such changes in VOT values of the L1 have also been reported in several other studies (e.g. Caramazza et al., 1973; Williams, 1979).

Similar results were found by Beckman (1986) in a study on the perceptual cues to accent in English and Japanese. Three groups of subjects were tested on the effects of fundamental frequency (F0), amplitude, spectral quality and duration patterns on the perception of accent in English and Japanese. The three groups were Japanese, American speakers of Japanese, and American monolinguals. The results

---

Several prosodic factors have been recognised, among which intonation, rhythm, stress, loudness and tempo (Crystal, 1995; Ladefoged, 1993).

<sup>6</sup> Voice onset time (VOT) is defined "as the duration of the interval between the consonant opening (in the oral part of the [vocal] tract) and the onset of voicing at the larynx" (Lieberman, 1996). VOT is an important variable in the distinction between voiced and voiceless stops. Abramson and Lisker (1970), in a study on cross-linguistic differences in the perception of VOT by English, Spanish and Thai speakers, were the first to establish how the acoustic boundaries for the voicing distinction vary with different languages.

of her experiment are shown in Figure 1.1. Confronted with Japanese stimuli, the mean effect-on-accent scores for the fundamental frequency were higher for Japanese subjects than for the American speakers of Japanese, who in turn scored higher than the American monolinguals. In other words, when listening to Japanese stimuli, American speakers of Japanese used F0 cues to an extent which was intermediate to that of Japanese and monolingual English subjects.

The results for the English stimuli showed that Japanese subjects used F0 to a much greater extent than the other cues to accent. The American monolinguals, when confronted with the same English stimuli, used all four cues to a similar extent. The American speakers of Japanese, on the other hand, seemed to use F0 more than the American monolinguals did, although they did not rely as much on this cue as the Japanese subjects. The spectral and durational patterns showed a score intermediate between the scores of the Japanese and monolingual American subjects.

The Japanese subjects in this experiment were not monolinguals, and had been exposed to English for a considerable amount of time (some of them had been living in the States for 10 years or more). The fact that they kept using F0 as their main cue to accent suggests that it is difficult for Japanese to learn to use other cues to accent. The American speakers of Japanese, on the other hand, seem to have less difficulty in learning to use the F0 cue more than the other cues. Apparently, it is relatively easy for American speakers to learn to use F0 more as a cue to accent, whereas it is more difficult for Japanese to learn to use other cues than pitch. This suggests that the fundamental frequency cue may be a universal and unmarked feature, and therefore easier to access for L2 learners (for a similar conclusion see Kondo, 1995).

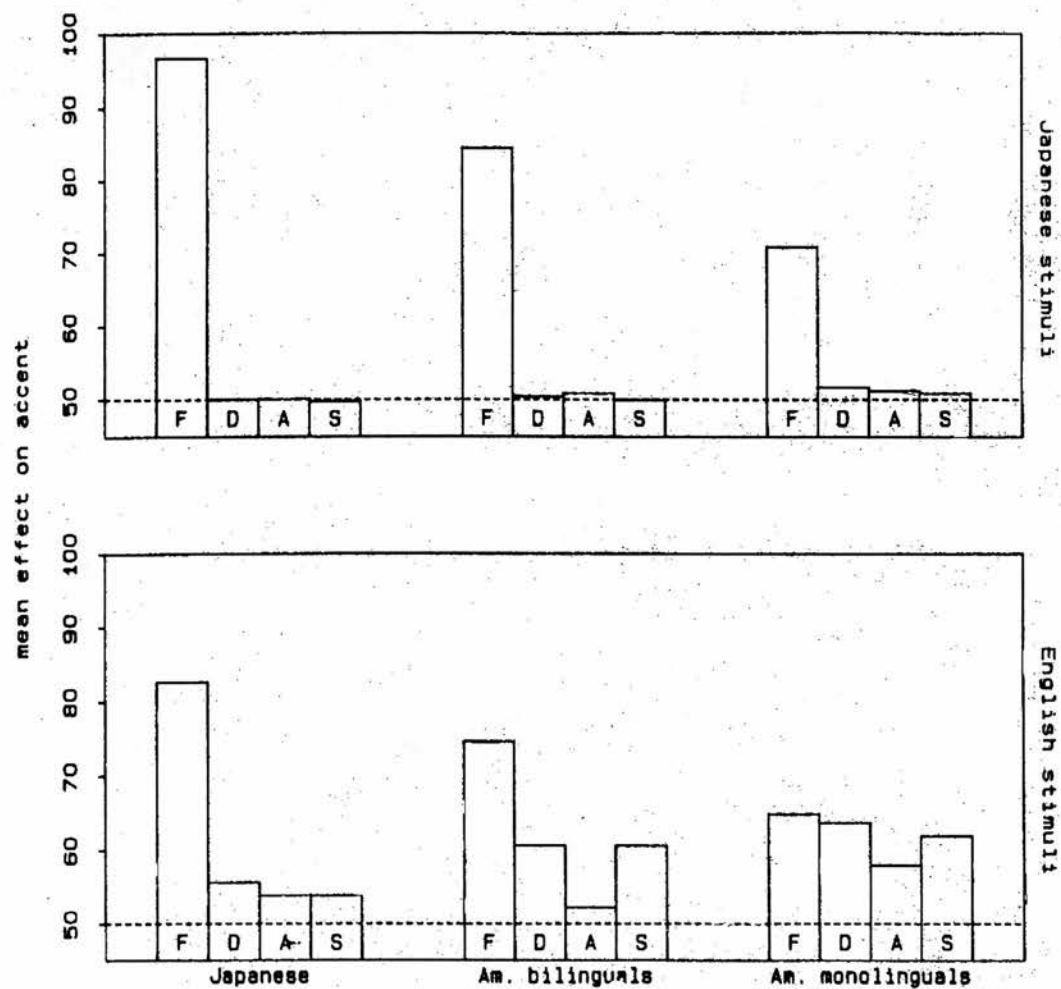


Figure 1.1. Mean effect on accent of the fundamental frequency (F), duration (D), amplitude (A), and spectral coefficient (S) patterns, averaged over responses to all tokens of Japanese stimuli (upper graph) and English stimuli (lower graph). Means are broken down by subject group. (adapted from Beckman, 1986).

### 3. Factors that affect speech production and perception in L2

With a substantial number of studies carried out in the last two decades, new facts have been brought to light in the area of interlanguage phonology. In this section I will present the most important factors which are thought to affect L2 speech production and perception. It should be noted that the factors discussed in this section are not the only ones that can affect L2 speech production and perception. A number of other factors (i.e. motivation, sex, personality variables, social distance) may also play a role in the learner's success in the acquisition of L2 speech. However, these non-linguistic factors will not be discussed in this thesis (but see, for example Leather and James, 1996).

#### 3.1. Native language influence

The fact that the native language almost invariably influences the perception and production of sounds in a second language has been widely recognized (cf. Jenkins and Yeni-Komshian, 1995). Researchers mostly agree on the role of language transfer in second language acquisition, although it is believed that other factors can be involved as well (as is described extensively in section 2.2). It is not entirely clear what exactly in the native language contributes to the difficulty in acquiring second language speech. Some scholars assume that it is a perceptual difficulty which causes someone to have a foreign accent when speaking a foreign language. They believe that non-native speakers often do not perceive the L2 sounds in exactly the same way as do monolingual speakers of the L2. This assumption goes back a long way, to the idea that the L1 sound system acts as a 'phonological filter' (Trubetzkoy, 1939) or a 'grid' (Wode, 1980) through which the L2 sounds are perceived. Even as early as 1931, scholars (Polivanov, 1931) assumed that errors in the

production of speech sounds resulted from inappropriate use of previously acquired L1 phonological structures. Some researchers reporting on the second language acquisition of prosody even suggested that L2 learners are tone-deaf (Buysschaert, 1990) or phonologically deaf (Lepetit, 1989), a perceptual difficulty which learners are unlikely to overcome. The idea that accented pronunciation of L2 sounds may be perceptually motivated, is taken over in recent studies (Best and Strange, 1992; Flege, 1995<sup>7</sup>; Rochet, 1995; Rvachew and Jamieson, 1995). For example, Flege (1981;1987a;1987b) believes that a foreign accent results from the development of 'inaccurate perceptual targets'. Flege (1981;1987a;1987b) introduced the concept of '*equivalence classification*' as a cause for the persistent foreign accent of many adult L2 learners. Equivalence classification is a cognitive mechanism which causes learners to classify similar sounds in L1 and L2 into a single category. When the sounds in L1 and L2 are 'similar', the learners cannot develop a new perceptual target, but place the L2 sound in the same category as the 'similar' L1 sound. As a result, they develop inaccurate perceptual targets for L2 sounds. More discussion on this will follow shortly.

However, it has been convincingly demonstrated that in order to produce a L2 contrast, it is not always necessary to be able to perceive this contrast accurately. For example, Goto (1971) and Sheldon and Strange (1982) reported that their Japanese subjects produced the English /r/ and /l/ contrast more accurately than they were able to perceive it. So the assumption that if L2 learners can produce a phonetic distinction, they must be able to perceive it, is not true. Recent work by Yamada *et al.* (1995) showed however, that improved perception can have a positive effect on production. Yamada *et al.* trained Japanese speakers in the English /r/-/l/ contrast, and found that at the end of the training they were able to produce this contrast significantly better. Although the relation between

---

<sup>7</sup> Although Flege (1995) believes that many production errors in L2 speech result from perceptual difficulties, he is not convinced that it can explain all production errors (p.

perception and production is still not entirely clear, this finding suggests that foreign accents at least partially result from perceptual difficulties.

### 3.2. Types of phonetic categories and contrasts

Not all second language phonetic categories and contrasts constitute the same amount of difficulty for L2 learners. Several sources for this phenomenon have been reported. One source appears to be the nature of the acoustic signal. In a recent review on speech perception in L2 acquisition, Jenkins and Yeni-Komshian (1995) report that a considerable amount of past research suggests that it is relatively easy to master the perception of temporal aspects of L2 speech phenomena, such as VOT. On the other hand, they suggest that modification of the perception of phenomena like spectral change or place contrasts appear to be much more difficult for L2 learners.

Another source of the fact that some phonetic contrasts appear to be more difficult than others lies in the degree of similarity between the learner's L1 and the L2. At the time of the contrastive analysis approach, it was thought that 'new' second language sounds, i.e. sounds which do not exist in the learners' L1, are difficult for learners. The prediction was that, for example, the French vowel /y/ would be difficult to acquire for English speakers, since the high front unrounded vowel /y/ does not exist in English. French /u/, on the other hand, should not constitute a problem for English speakers, as it does have a counterpart in English. This view proved to be inadequate. Acoustic measurements (Flege, 1987a; Flege, 1987c) revealed that the opposite was true, American English learners of French pronounced French /y/ more accurately than /u/. Recently, two models (Best, 1995; Flege, 1995) have been developed to account for and predict the degree of difficulty of L2 sounds. In the above case, both Flege and Best would say that when two sounds in L1 and L2

are 'similar', like French and English /u/, learners will not modify their perception or production of these sounds. Perception will be relatively easy, as the L2 sound will be placed in the same category as the 'similar' L1 sound. Production, on the other hand, will be accented, as the learner does not modify his L1 production for a 'similar' L2 sound. However, in the above case of French /y/, the L2 sound is noticeably different from any phonetic category in L1. Here, Flege would argue that learners will develop a new category for this 'new' L2 sound, and Best would argue that perception should not be a problem. The models of Flege and Best will be discussed in more detail below.

### 3.3. Age of learning

It has been reported in a number of studies that there is a relation between the age of learning and the degree of 'foreign accent' in the production of a second language (c.f. Flege, Munro, and Mackay, 1995; Long, 1990). The notion of a critical period for language learning was introduced by Lenneberg (1967) who observed that "foreign accents [in a second language] cannot be overcome easily after puberty" (p. 176). In a review of age-related effects on L2 acquisition, Long (1990), concluded that there are several 'sensitive periods'. Each linguistic domain has its own sensitive period, after which acquisition is irregular and incomplete, with that for phonology being the earliest (age 6) and syntax the latest (age 15). He concluded that learners will produce an L2 without foreign accent if learning starts before the age of 6. If learning begins after the age of 12, a foreign accent will be the result, and starting between the age of 6 and 12 will lead to variable success. It has also been suggested, that the sensitive period does not end abruptly (Long, 1990), but that there is a linear relation between age of learning and the degree of perceived foreign accent (Flege, Munro, and Mackay, 1995). It is possible for individuals to have a foreign accent in their L2, although learning started in childhood

(Long, 1990). Flege and colleagues even reported an age of 3.2 years for the onset of effects on the L2 pronunciation (Flege, Frieda, and Nozawa, 1997).

In other words, the earlier in life one begins to learn a second language, the better one is able to pronounce it. Flege, Munro and MacKay (1997) summarize it as follows.

Both the *proportion* of individuals observed to speak their L2 with a detectable accent, as well as the *strength of perceived foreign accents* among individuals with detectable foreign accents have been found to increase as the age of learning an L2 increases (Flege *et al.* 1997: p. 3125).

### 3.4. Language experience and ultimate attainment

A substantial number of studies conducted in the 1980s and 1990s demonstrate that increased exposure to the second language may improve the production and perception of L2 speech (e.g. Best and Strange, 1992; Bohn and Flege, 1990; Flege, 1987b; Flege, 1995; Oyama, 1976; Sheldon and Strange, 1982; Strange, 1995a). A non-native speaker of Greek who has lived in Greece for more than ten years is likely to be more proficient than a person who has only been exposed to Greek for a couple of months. However, language exposure is not a guarantee for an accent-free L2 pronunciation. As we have seen before (section 2.2 and 2.3), some researchers believe that it is *never* possible for adult L2 learners to speak a second language without a trace of foreign accent, or at least not without it having an effect on the L1. Major (1990) for example, suggests that it is impossible for adult learners to maintain native-like pronunciation in both L1 and L2. He based this assumption on evidence for VOT values of voiceless stop consonants, obtained from experienced L2 learners. Apparently, the more closely the VOT values of the learners resembled the L2 norm, the more their L1 stops started to resemble L2 stops. He concluded that L2 learners can either (i) fail to achieve accent-free L2 speech and maintain native L1 pronunciation; (ii) achieve native-like L2 pronunciation but lose native L1 pronunciation; or (iii) lose

native L1 pronunciation but still fail to achieve native-like L2 pronunciation. Proficiency in both L1 and L2 pronunciation, he claims, is not possible, because of the mutual effects of L1 and L2. This suggests that he assumes that the L1 and L2 phonetic systems are not (fully) isolated, which causes them to interact with one another. This bidirectional interference, L1 influencing L2 and vice versa, has been attested in some studies. For example, the results obtained by Flege and Hillenbrand (1984) and Beckman (1986) described in detail in section 2.3, seem to support this. In the latest adaptation of his Speech Learning Model (SLM), Flege (1995) states that the model does not predict full mastery of certain L2 sounds anymore, and that it agrees with Grosjean's (1989) view on bilingualism, in which a bilingual's two language systems are always activated to some degree. "The model postulates that L1 and L2 sounds exist in a common phonological space" (Flege, Munro, and Mackay, 1995: p. 3133) This implies that the L1 and L2 necessarily interact, and therefore L2 learners will never reach full competence in L2.

Nevertheless, some studies presented by Neufeld, indicate that full mastery of L2 pronunciation and intonation *is* attainable in adulthood (Neufeld, 1977; 1979; 1987)<sup>8</sup>. His studies, report on the exceptionally high levels of pronunciation and intonation L2 learners can achieve after short intensive instruction. His studies are however strongly criticized by many researchers who question their methodology (Long, 1990; Scovel, 1981; Tarone, 1987).

Whether accent-free pronunciation of L2 is attainable or not, there is an abundance of evidence that there are considerable differences in the proficiency of language learners with the same amount of language experience (e.g., McKain, Best and Strange 1981, Beckman, 1986; Flege, 1992; Hecht & Mulford, 1987; Jenkins & Yeni-Komshian, 1995; Yamada *et al.*, 1995). Almost any study on L2 speech finds outliers, like learners with

---

<sup>8</sup> Williams (1980) also presented evidence that some adults appear capable of producing a foreign language without accent.

excellent pronunciation skills, despite very limited exposure to L2, and other learners who even after many years of exposure do not seem to have improved at all. Some researchers wonder if some speakers have a 'gift' for languages, whereas others don't. Cases of superexceptional talent for L2 learning have been reported in a number of studies (Ioup, 1995; Ioup et al., 1994; Novoa, Fein, and Obler, 1988; Schneiderman and Desmarais, 1988). But the source of individual differences is still unclear. At the moment there is no satisfactory explanation for these sometimes striking individual differences in L2 speech production and perception (Flege, 1992). It is not clear if large individual differences also exist within first language acquisition (Jenkins and Yeni-Komshian, 1995). Hecht and Mulford (1987) posit that the individual differences found within first language acquisition, make a comparison of the development of phonology in L1 and L2 learners very complicated. They state that the variability found in L2 speech, might well be attributable to "the same forces that produce variable pronunciations during the course of first language development, for example spontaneous production versus imitation, conflicting or changing phonological rules, and the influence of particular lexical items" (p. 227).

In order to come to any conclusive finding about whether it is possible for adult learners to ultimately attain a nativelike accent in an L2, it is necessary to include very advanced and successful late learners in L2 studies. However, most studies that address age-related differences in L2 acquisition have not been specifically designed to address the issue of ultimate attainment. To my knowledge there are only a few studies that address this issue, of which some tested syntactic properties of L2 acquisition (Birdsong, 1992; Coppieters, 1987; Sorace, 1993), others tested ultimate attainment in L2 pronunciation (Bongaerts, 1997; Bongaerts, Planken, and Schils, 1995; Bongaerts et al., 1997). Most of these studies suggest that native-like performance or competence in the L2 is possible, albeit rather exceptional, in the aspects tested. However, native-like competence or performance in one aspect of language does not necessarily

imply native-like performance or competence in other aspects (Sorace, 1993).

### 3.5. Effect of training

Recent perceptual training studies carried out in the last decade, have indicated that intensive training can improve the perception abilities of non-native contrasts. Although some contrasts are easier than others, perception of all contrasts seem to benefit from training (Rochet, 1995; Rvachew and Jamieson, 1995). Nevertheless, generalisation to other contexts, like new speakers, new words, or different syllable positions, appears to be limited (Strange, 1995b). Other research suggests that perceptual training can sometimes improve production, in other words, production might benefit from perceptual training.<sup>9</sup> But the effect of perceptual training on the production of speech sounds, has not been studied extensively. Furthermore, some studies suggest that articulatory instruction or training can also lead to improved production and perception (Catford and Pisoni, 1970; Greasly, 1971; Weiss, 1976). Taken together, the findings do not suggest a constant and simple relationship between perception and production.

---

<sup>9</sup>Yamada *et al* (1995) found that the production of the English /r/-/l/ contrast by Japanese speakers, significantly improved after perceptual training of this contrast. Other researchers also found that an improvement in perceptual performance (after intensive perceptual training) carried over to production. For example, Rochet (1995) found that perceptual training of the voicing contrast in French stop consonants, resulted in an improved production of these contrasts by adult Mandarin Chinese speakers. Rvachew and Jamieson (1995) reported a similar effect on the speech production of children with speech disorders.

#### 4. Current second language speech models

Recently, two working models have been proposed that try to account for the difficulties L2 learners have in perceiving or producing the sounds of a second language<sup>10</sup>, Flege's (Flege, 1995; Flege, Munro, and Mackay, 1995) Speech Learning Model (SLM) and Best's (Best and Strange, 1992; Best, 1995) Perceptual Assimilation Model (PAM).

Both models are concerned with the relation between L1 and L2 sounds, and they both describe this relation in terms of 'phonetic similarity' between the two languages under consideration. Flege's SLM describes this similarity in *acoustic-phonetic* terms. The degree of perceived similarity between the L1 and L2 sounds will determine whether new categories for L2 sounds can be established. When the L2 sound is noticeably different from the L1 sound, the learner will eventually develop a new category for this 'new' L2 sound. Because a new category is established, the sound can be produced in an accent-free manner, although there is no guarantee that this will actually happen (as will be explained below). When, on the other hand, L1 and L2 sounds are acoustically 'similar' (but not identical), the cognitive mechanism of equivalence classification (cf. section 3.1) causes the listener to classify them into a single category. This is because during L1 acquisition we have become attuned to certain phonetic contrasts, and we fail to recognize that the differences between L1 and L2 sounds are phonetically relevant. As a result, the formation of a new category is blocked. When a category cannot be established for an L2 sound, production of this sound will be phonetically inaccurate, resulting in accented production. It is also predicted that the production of the 'similar' L1 sound will gradually shift

---

<sup>10</sup>Both models only account for *segmental* aspects of foreign accent, although they do state that nonsegmental dimensions are an important source of foreign accent (e.g. Best, 1995: p. 192; Flege, 1995: p.233; Flege *et al.*, 1995).

away from the monolingual norm. In other words, the L2 sound and its L1 counterpart will gradually 'merge' (Flege and Hillenbrand, 1984).

Best's PAM, like Flege's SLM, also recognizes the similarity between L1 and L2 sounds as a source of difficulty in the perception of L2 sounds. Unlike Flege, Best defines the similarity between L1 and L2 sounds in *articulatory-phonetic* (gestural) terms. The PAM predicts that L2 phonetic segments and contrasts<sup>11</sup> are assimilated to L1 categories, according to their degree of similarity to L1 gestural constellations. Different assimilation patterns are possible, and each different pattern predicts the degree of difficulty in the perception of an L2 contrast. The following assimilation patterns for non-native contrasts are possible: (i) if both contrasting L2 sounds are very discrepant from any L1 phonetic gestures, they either cannot be categorized into an L1 category, or they are categorized as non-speech sounds. Perceptual differentiation between the two pairs of the contrast is thought to be good; (ii) perceptual differentiation will be equally good (if not better), if the two sounds are assimilated into *two different L1 categories*; However, (iii) perception will be poor, if both pairs of the L2 contrast are assimilated into a *single L1 category*; (iv) perception will be slightly easier, if both members of the L2 category are assimilated to a single L1 category, but one is seen as more deviant from the native 'ideal' than the other (they differ in their *category goodness* of fit to the L1 category).

Although the two models seem quite different at first sight, many of their assumptions are similar. Both models assume that an L2 sound which is totally different from any phonetic category in the L1 should not constitute any difficulty to L2 learners. But Best is mainly concerned with the perceptual difficulties of the L2 learner, whereas Flege tries to account for their production difficulties. In the above case, Best would argue that the perception would be good to excellent, whereas Flege would argue

---

<sup>11</sup> Best and Strange (1992) noted that the PAM and SLM differ in this respect, since the former is mainly concerned with the perception of L2 *contrasts*, whereas the latter is more

that the learner will develop a new target for this sound, and accent-free production of this 'new' sound is possible. The two models also agree on the fact that sounds which are similar but not identical to L1 sounds, are the most problematic for L2 learners. Flege's model assumes that 'similar' sounds get equated with an L1 sound, and they continue to be classified to L1 categories, even after long exposure. This, according to Flege, results in 'accented' production. Best assumes that learners are able to perceive variations in the goodness of fit of an L2 sound to an L1 category.

As mentioned before, when the L1 and L2 sounds are different, both models assume that perception and production of this new L2 sound should be unproblematic. But, this is not always the case. The SLM assumes that L1 and L2 sounds exist in a common phonological space.<sup>12</sup> Flege hypothesises that in order to maintain contrast within that common space, that is contrast *within* and *across* languages, the categories might be 'deflected away' from each other (Flege, 1995: p. 242; Flege, Munro, and Mackay, 1995: p. 3133). As a result, the L2 sound might not be produced in the same way as it is produced by native speakers. In fact, when speakers manage to maintain phonetic contrast between the categories in the L1 *and* L2, the contrast is likely to be inaccurately produced in *both* languages.<sup>13</sup> According to the PAM, the perception of an L2 category which is different from any L1 category, can sometimes be poor when it is close to native categories (Best, 1995).

Although both models depend on the notion of phonetic similarity, they differ in the way they define it. Flege, for example, focuses on *acoustic-phonetic* similarity, whereas Best defines the similarity in *articulatory-phonetic* (gestural) terms. In fact, both authors believe that the two are not easy to separate, and that both kind of similarities play a

---

concerned with *segments* .

<sup>12</sup> The PAM does not (as yet) take a stance regarding the issue of a common or separate phonological space (Best, 1995: p.198).

<sup>13</sup>By inaccurate production Flege refers to slight differences in production compared to the

role in L2 sound acquisition (Best & Strange, 1992: p. 306; Flege & Hillenbrand, 1984: p. 708). It is obvious, as both Best and Flege admit (cf. Bohn, 1995: p. 90; Strange, 1995b: p. 81), that the notion of phonetic similarity needs to be more precisely defined. Objective criteria need to be established for deciding which sounds are 'new' and which are 'similar'. But for the moment, the two models can be used as working models, both useful in different ways. Currently, Best's model seems to be more useful for the perception of sounds that can be distinguished by articulatory cues, such as consonants. The SLM seems to be more suitable for those sounds that cannot be easily distinguished on articulatory terms, like vowel height and backness. Also, the PAM is more concerned with the perception of L2 categories, whereas the SLM is intended to address perceived similarities between individual segments in L1 and L2. From this distinction one might assume that the PAM is better equipped, for instance, to account for perception difficulties of Japanese learners of the English r-l contrast, a contrast which does not exist in Japanese. The SLM is probably more suitable to account for subtle acoustic differences, like slight deviations from the native norm in the pronunciation of English /t/ by French learners of English.

Although both Best's and Flege's models, like the earlier contrastive analysis, assume that adult learners tend to interpret L2 sounds in terms of their L1 systems, the predictions made by the current theories are very different from those made by the contrastive analysis (see also Section 2.1). The contrastive analysis would predict that the English /t/ would not constitute a problem for native speakers of French, because English and French /t/ are fairly similar. The French learners would simply use the French /t/ for both English and French. The SLM, on the other hand, would predict that just because the English and French /t/ are rather similar (but not identical), the production of the French /t/ would constitute a problem for French learners. The PAM cannot make such

---

native norm. They might not be detectable by ear, but only by acoustic measurements.

predictions, as it is concerned with the degree of difficulty in the perception of an L2 contrast, like the /t/ - /d/ contrast. However, just like the SLM and contrary to the contrastive analysis, the PAM assumes that sounds which are similar to L1 sounds, are the most problematic for L2 learners.

This difference in predictions between the contrastive analysis and the more recent speech models, may be due to methodological differences between these approaches. For example, contrastive analysis was based on impressionistic observations, mostly done in language classrooms. The more recent models (SLM and PAM) had the advantage of developments in speech technology, and are based on instrumental data which are analysed in speech laboratories. It is obvious that, when judging L2 speech auditorily, substitution of an English /t/ by a French /t/ is not as striking as, for example, the substitution of a Dutch /d/ for the English /ð/.

Furthermore, the production data on which Lado's contrastive analysis is based were gathered in language classes. This implies that the contrastive analysis is based on language that is learned formally outside the language community, and presumably the learners were not at the most advanced stages of language learning. Flege's model on the other hand is concerned primarily with the ultimate attainment of L2 pronunciation (Flege, 1995). His model is based on production data of advanced learners, who learned the L2 within the language community. According to Flege's model, learners are not immediately able to discern differences between L1 and L2 sounds, even when the L2 sound is 'new'. So, the beginning learner will have difficulties with the production of both 'new' and 'similar' sounds. And the difficulty the beginning L2 learner will have producing such a 'new' sound must surely be much more striking than the substitution of a 'similar' L2 sound by the closest L1 sound. Therefore, the discrepancy between Flege's and Lado's predictions may, to a certain extent, be due to the level of learners reported in their work.

#### 4.1. Criteria for measuring phonetic similarity

In the previous section it was seen that both the PAM and the SLM base their predictions of which L2 sounds (or contrasts) will be difficult for L2 learners on the degree of similarity between L1 and L2 sounds. However, objective criteria for measuring this degree of similarity remain to be established (Bohn, 1995: p. 90; Leather and James, 1996: p. 276; Strange, 1995c: p. 81).

The closest the SLM has come to defining similarity between L1 and L2 sounds is the assumption that L1 and L2 sounds that are transcribed by the same IPA (International Phonetic Alphabet) symbol can be seen as 'similar' (Flege, 1987a; Flege and Hillenbrand, 1984). The following example of similarity is given:

... instances of /t/ occurring in French and English words are likely to be regarded by the L2 learner as being different realizations of the same category because of their overall phonetic similarity (Flege and Hillenbrand, 1984: p. 708).

However, this does not explain whether, for example, French or Arabic speakers of English would differ in their ability to produce English /t/, since these languages vary in phonetic detail of the sound /t/. The VOT values for syllable-initial /t/ vary from 20 ms for the French /t/, 37 ms for Arabic /t/, to 75 ms for English /t/ (for French: Caramazza *et al.*, 1973; for Arabic: Flege, 1980; for English: Flege, 1987a).

Best's PAM defines similarity in terms of similar gestural constellations of L1 and L2 categories. She gives the following examples

For a native listener of a language that has no dental stop but does have bilabial, alveolar, and velar stops, the tongue tip constriction of the dental stop is straightforwardly closer in native phonological space to the alveolar place than to the others, because the articulation involved is the same and the place of constriction is more similar than those of bilabial or velar stops (Strange, 1995: p. 193-4).

However, this does not predict why, for example, French speakers substitute /s/ and /z/ for English /θ/ and /ð/, whereas Russian speakers prefer /t/ and /d/ for the same phonemes, when both French and Russian have /t d s z/ in their L1 phoneme inventories (Weinberger, 1990).

Another shortcoming of the SLM and PAM is the fact that these models are developed to account for *segmental* aspects of L2 acquisition only. So far, no attempt has been made to use these models to account for prosodic aspects of L2 acquisition. One of the aims of this thesis is to test whether one of these models (Flege's SLM) is also able to account for prosodic aspects of L2 learning (in this case intonation). Only the SLM was tested as it was thought that this model can better account for the *production* data examined in this thesis than the PAM, which is only concerned with the perception of L2 speech (as described in section 4). In Chapter 2 it will be shown how Flege's model can be applied to L2 intonation data.

## 5. Second language acquisition of intonation

### 5.1. An overview

As we have seen in the previous sections, most of the studies on interlanguage phonology have dealt with the acquisition of L2 segments only. The prosodic (suprasegmental) aspect of interlanguage phonology has been largely ignored. But just as languages differ in their inventory of phonemically contrasted consonants and vowels, languages also differ in their prosodic patterns. For example, languages can be different in their inventory of distinct melodies, or in the manner that prominent words are differentiated from less prominent words. Furthermore, just as the exact phonetic realisation of segments can differ across languages, the phonetic realisation of phonologically distinct melodies might differ as

well. Thus, as with phonetic segments, the prosodic characteristics of a speaker's native language are likely to interfere with the production and perception of prosodic aspects of a second language.

It is generally accepted that L2 learners often make prosodic errors and that these errors can be a cue to foreign accent (Flege, 1992; Flege, 1995; Flege, Munro, and Mackay, 1995; Magen, 1998; Munro, 1995; Willems, 1982). Nevertheless, not many studies have dealt with the acquisition of suprasegmental characteristics, like intonation, stress, or rhythm, of a second language. The absence of studies on the prosodic patterns in an L2 has been pointed out by several researchers, who argue that we need to start looking into this aspect of L2 learning (Anderson-Hsieh, Johnson, and Koehler, 1992; De Bot, 1986; Jenkins and Yeni-Komshian, 1995; Munro, 1995).

Lately, interest in the role that prosodic phenomena play in identifying a speaker as non-native has been growing. Munro (1995) was able to show convincingly that it is possible to detect foreign accent on the basis of nonsegmental information alone. Untrained listeners were presented with English sentences and a narrative, produced by Mandarin and English speakers, of which most of the segmental information was filtered out. On the basis of the remaining prosodic information, like fundamental frequency, word durations and rhythmic properties, listeners could accurately determine whether an utterance was made by a native or non-native speaker of English. Munro also attempted to identify which of the properties of the prosodic information contributed to a perceived foreign accent. Several possible candidates were suggested, such as speaking rate, fundamental frequency patterns, and patterns of reduction (lack of reducing full [t] to a flap [ɾ]), but their exact contribution needs further research. A study by Van Els and De Bot (1987) also pointed towards the importance of intonation patterns as a cue to foreign accent.

The precise role that prosodic errors play in the perception of accented speech is not well understood. Some studies suggest that errors in prosody contribute more to perceived foreign accent than segmental

errors (Anderson-Hsieh, Johnson, and Koehler, 1992; Johansson, 1978), whereas others suggest the exact opposite (Flege, 1992). Considerably more research is needed to resolve this disagreement.

Several researchers have attempted to identify the prosodic errors L2 learners often make. The prosodic factor that is mostly reported to be affected in the L2 is intonation. Specifically, many studies report transfer of L1 intonation patterns to the L2 (Adams and Munro, 1978; Backman, 1979; Buysschaert, 1990; De Bot, 1986; Grover, Jamieson, and Dobrovolsky, 1987; Jenner, 1976; Willems, 1982). Other reported errors are differences between native and nonnative speakers in rhythm (Wenk, 1985), stress placement (Adams and Munro, 1978; Archibald, 1992), and the relationship between word stress and vowel reduction (Flege and Bohn, 1989). Studies on the intonation of L2 learners are mainly production studies, although one thorough study deals with perception difficulties. In this study, Cruz-Ferreira (1983) presented two groups of L2 learners (Portuguese learning English, and English learning Portuguese) with a set of minimally paired sentences in each L2, differing in intonation only. Listeners were asked to decide whether the sentences in each pair had the same or a different meaning. They also had to match the sentences with a meaning gloss. It was found that both groups of nonnatives interpreted the sentences correctly, when the intonation patterns in L1 and L2 were similar, or when a 'universal' intonation pattern was used. In the former case, the nonnative listener employs a strategy of positive transfer, whereas in the latter case a 'pitch height strategy' is used, referring to the listener's use of 'general intuitions about the more likely meanings associated with lower and higher pitch' (Cruz-Ferreira, 1987: p. 116). On the other hand, sentences were incorrectly interpreted (negative transfer), when the intonation patterns in L1 and L2 were the same, but the use of the pattern is different. Listeners also made errors when the sentence contained lexical or grammatical items which are usually associated with an 'unmarked' meaning in the L2. In this case the 'lexico-syntactic strategy' was used, the more unusual 'marked' meaning was rejected, and

the more straightforward meaning was assigned to the intonation pattern.

Backman (1979) also raised the question of universals in acquiring the intonational system of a second language, but contrary to Cruz-Ferreira (1987) she based this on a production study. She observed that the errors she found in her study of the English of Spanish learners showed remarkable similarities with errors Jenner (1976) found in his study on the English of Dutch learners. More recent studies support the finding of a similarity in errors in the production of L2 English intonation by speakers with different language backgrounds. The errors reported were (i) a narrower pitch range (Backman, 1979; Jenner, 1976; Willems, 1982); (ii) problems with the correct placement of prominence (Backman, 1979; Jenner, 1976); (iii) replacement of rises with falls and vice versa (Adams and Munro, 1978; Backman, 1979; Jenner, 1976; Lepetit, 1989; Willems, 1982); (iv) incorrect pitch on unstressed syllables (Backman, 1979: too high; Willems, 1982: no gradual rise on unaccented words preceding a fall); (v) difference in final pitch rise (Backman, 1979: too low; Willems, 1982: too high (overshoot)); (vi) starting pitch too low (Backman, 1979; Willems, 1982); (vii) problems with reset from low level to mid level after a boundary (Willems, 1982); and (viii) a smaller declination rate (Willems, 1982). Although it is true that some of the observed errors are similar, it should be emphasized that they all appeared in studies on English as a second language. So the similarities might be due to idiosyncrasies of the English intonational system. Furthermore, the similarities cannot be explained by developmental factors alone. For example, the fact that both Dutch and Spanish acquiring English intonation produce a smaller pitch range compared to native English speakers does not necessarily indicate that a reduction of pitch range is a universal tendency in L2 acquisition. The smaller pitch range in the data of the learners could simply be a case of transfer, since both Dutch (Jenner, 1976) and Spanish (Stockwell and Bowen, 1965) are reported to have a smaller pitch range than English. It is therefore more likely that there is

more than one process involved in the acquisition of L2 intonation, a conclusion which has also been reached in other fields of L2 acquisition (see also section 2.2).

## 5.2. Describing intonation across languages

It should be noted that comparison of the findings described in the previous section is not an easy task. The studies differ considerably with respect to the proficiency level of the learners, the languages under investigation, and the framework or methodology used in the study. These differences in methodology prevent us from coming to any reliable conclusions about the similarities and differences between the languages investigated in these studies, and the process of L2 acquisition of intonation.

One of the major problems in comparing these studies is the fact that they are to a great extent based on impressionistic descriptions of intonation, such as those by Halliday (1970) and O'Connor and Arnold (1973). These pedagogical courses were admirable attempts to give a full account of English intonation and need to be valued as such. However, due to the lack of technical facilities at the time they were written, they are essentially based on impressionistic observations. This means that instrumental evidence for their claims is lacking. It is exactly this lack of instrumental evidence which makes comparisons across languages and between native and non-native speakers of a language extremely difficult.

Without instrumental evidence it may be difficult to characterise some cross-linguistic differences which are not easily detectable by ear. For example, it has been shown in a number of cross-linguistic and second language acquisition studies that some phonetic properties exhibit language-specific characteristics. Phonetic properties which are found to differ across languages are, for example, coarticulatory strategies (Hardcastle, 1982; Kondo, 1995), the degree to which vowels are more

central or peripheral in the vowel space (Flege, 1984; Flege and Hillenbrand, 1984; Jongman, Fourakis, and Sereno, 1989) and the voice onset times (VOT) for voiceless segments (Caramazza et al., 1973; Flege, 1981; Flege, 1984; Willems, 1982). These fine cross-language phonetic variations are not phonologically contrastive, and may be transcribed by the same IPA symbol. Nevertheless, they seem to contribute to the characteristic sound patterns of languages (Flege, 1984; Flege & Hillenbrand, 1984). It may therefore be important to pay attention to this kind of phonetic detail when comparing languages, or when investigating aspects of second language acquisition of speech.

This thesis will focus on phonetic aspects of L2 intonation. However, in order to explain some of these phonetic aspects, some phonological aspects of L2 intonation need to be taken into account. In the next sections, I will give a description of some (phonetic and phonological) properties of intonation which are likely to be affected in L2 speech production.

Before these properties of L2 intonation are presented, the theoretical assumptions adopted in this thesis will be described. This thesis is not aimed at providing a model for the description of L2 intonation, nor at resolving issues in phonological theory. Its aim is rather to provide a fairly theory-independent description and account of the data on L2 intonation. However, some basic notions which will be used throughout this thesis are adopted from some of the research based on the autosegmental-metrical approach to the study of intonation (e.g. Pierrehumbert, 1980; Pierrehumbert and Beckman, 1988). In this approach the existence of two types of intonational events is recognised: pitch accents and boundary phenomena. *Pitch accents* are defined as distinctive pitch movements associated with stressed or prominent syllables. *Boundary phenomena* are peripheral events, i.e. they are phonologically associated with some prosodic boundary (such as the end of a phrase or utterance). However, they are not expected to show phonetic alignment with a specific tone-bearing unit, such as a stressed syllable

(Gussenhoven, in press; Pierrehumbert and Beckman, 1988). The boundary phenomena are usually further divided into *phrase accents* and *boundary tones*. Boundary tones are thought to occur at the edge of intonational phrases, whereas phrase accents are thought to occur at the edge of an intermediate phrase, or after the final pitch accent and before the boundary tone (Pierrehumbert and Beckman 1988).<sup>14</sup>

Since much work on Dutch intonation is based on the approach developed at the Institute for Perception Research (IPO), the IPO analyses of Dutch intonation will be given together with a 'translation' of their analyses into autosegmental-metrical terms. As will be shown in Chapter 2, some of the notions from the autosegmental-metrical approach (like pitch accents and boundary tones) have clear correspondances in the IPO approach.

### 5.3. Possible errors in L2 intonation

#### 5.3.1. Alignment

Alignment is the time of occurrence of a high (H) or a low (L) tone relative to the segmental string (i.e. the timing of a peak or valley with

---

<sup>14</sup>The idea of a 'phrase accent' appears to stem from Bruce, specifically from his analysis of Swedish word accents (Bruce 1977). However, he used the term 'sentence accent' for what became known as 'phrase accent' through the work of Pierrehumbert (1980). Recently, there is some disagreement as to the phonetic alignment of the phrase accent with the segmental string (i.e. the timing of the phrase accent with the vowels and consonants in speech). Although phrase accents and boundary tones are expected not to show phonetic alignment with specific tone-bearing units (TBU's), like stressed syllables (Pierrehumbert & Beckman 1988), some researchers have presented evidence that some phrase accents do associate with specific TBU's and that, despite their phonological association with boundaries, they are phonetically aligned in the same precise way pitch accents have been shown to align (Arvaniti, 1998; Ladd, Arvaniti, and Mennen, 1997; Gussenhoven, forthcoming).

the vowels and consonants in speech). Over the last decade or so, phonologists have become increasingly aware of the fact that the relation between pitch accents and the actual realisation of speech segments is fairly complicated (Arvaniti, Ladd, and Mennen, 1998; Caspers and Van Heuven, 1993; Prieto *et al.*, 1995; Rietveld and Gussenhoven, 1995; Silverman and Pierrehumbert, 1990). The timing of H or L tones can be influenced by a variety of factors. First of all, it can be influenced by categorical distinctions between e.g. a pitch accent and a boundary tone. That is, a pitch accent is thought to be associated with an accented syllable, whereas a boundary tone is said to associate with a boundary (e.g. Bruce, 1977; Pierrehumbert, 1980; Pierrehumbert and Beckman, 1988). However, the timing of tones can also be influenced by phonetic factors. Where the phonological categories define, in a rather nonspecific way, the coordination between an intonational contour and text, the phonetic factors make this association more specific. The phonology provides an association between, for example, a H\* tone and a prominent syllable, whereas phonetic implementation rules define whether this tone will be realised e.g., at the beginning or the end of that syllable.<sup>15</sup>

Rietveld and Gussenhoven (1995) found that in Dutch the nuclear fall is sensitive to the segmental composition of the accented syllable. The falling accent is aligned earlier when the syllable onset contains a sonorant rather than a stop, and when the onset is a cluster rather than a single consonant. It is aligned later, when the coda of the accented syllable is voiced. Prosodic effects on alignment were studied by Steele (1986), who found that the alignment of nuclear H\* accents in English varied with

---

<sup>15</sup> In Pierrehumbert's analysis of English intonation, the central tone of a pitch accent (i.e. the one which is associated with the accented syllable) is indicated with an asterisk, and often referred to as the 'starred tone' (e.g. Pierrehumbert, 1980; Pierrehumbert and Beckman, 1988). The construct of a 'starred tone' has been adopted in most work based on Pierrehumbert. Ladd (1983) proposed to make a distinction between association and alignment, where the former refers to an abstract phonological 'belonging together', and the latter refers to the phonetic temporal alignment with the segmental string.

syllable length (speech rate) and the number of following unaccented syllables. Silverman and Pierrehumbert (1990), building on Steele's findings, examined the alignment of prenuclear high (H\*) accent in English under a variety of prosodic conditions (speech rate, proximity of a following boundary, and number of following unaccented syllables). They found that the peak is aligned earlier if the next accented syllable follows without intervening syllables. In a study on the alignment of H\* accents in Mexican Spanish, Prieto *et al* (Prieto, Van Santen, and Hirschberg, 1995) found both segmental and prosodic effects. These studies show that the phonetic effects on alignment are very systematic and precise.

Apart from these influences of prosodic and segmental context on the timing of L and H tones, recent research has suggested that alignment exhibits certain language and dialect-specific characteristics, more or less like those found for voice onset time (Caramazza et al., 1973; Flege and Hillenbrand, 1984). That is, the same phonological category (i.e. the same phonological association) may be realised (aligned) differently in different languages or dialects. Differences in alignment have been found in cross-dialectal studies on Swedish (Bruce and Gårding, 1978) and Danish dialects (Grønnum, 1991), ethnic subvarieties of Singapore English (Lim, 1995), and varieties of British English (Nolan and Grabe, work in progress). Cross-linguistic differences in alignment have not been investigated extensively. However, Ladd (1996) suggests that such differences can be found when comparing the intonation of languages. He illustrates this with an example of a certain type of fall, which he describes as "a local peak associated with the accented syllable, followed by a rapid fall to low in the speaking range, followed by a more gradual fall to the end of the phrase or utterance" (Ladd 1996: p. 128). This fall can occur in Italian as well as in English (or German). However, its realisation is different in these two languages. Where the peak in English (or German) is rather late (at or near the end of the stressed syllable), it is early in Italian. The following rapid fall in English (or German) takes place

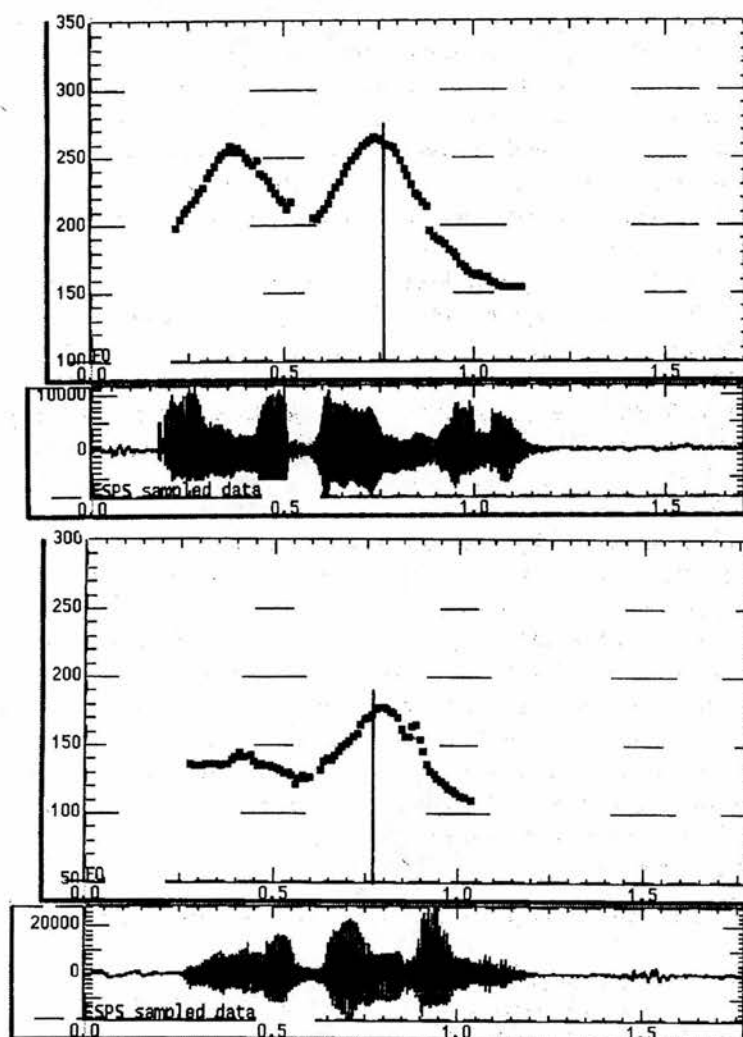


Figure 1.2. The Italian sentence 'E una vongola' (It's a clam) produced by a native Italian speaker (top panel) and by a non-native (English) speaker (bottom panel). The peak is reached earlier in the native speaker's rendition than in the English speaker's rendition. The vertical lines denote the boundary between the [ɔ] and [ŋ] of vongola. (adapted from Ladd, 1996)

between the stressed and following unstressed syllable, whereas in Italian the fall starts well before the following syllable.

As a consequence, English or German learners of Italian, may use their native alignment pattern when producing an Italian falling tune. In other words, the learner gets the association right, but fails to produce the correct alignment. Figure 1.2 gives an example of such a mistake. As Italians would place the fall near the beginning of the penultimate syllable, a delay of this fall may be interpreted by native Italians as a mistake in the placement of word stress, i.e. they may perceive this as stressed on the penultimate, rather than on the antepenultimate syllable.

It is for this reason that care needs to be taken when interpreting results on L2 intonation (especially when they are based on auditory observations only), which report errors in stress placement or replacement of rises with falls (e.g. Lepetit, 1989; Backman, 1979; Jenner, 1976). Some of these errors, may actually be phonetic errors (alignment errors), rather than phonological errors (misplaced stress). For example, Backman (1979), in her study on intonation errors of Venezuelan Spanish adult learners of American English, reports that the L2 learners often had problems with stress placement. However, visual inspection of some of the sample contours presented in her paper, suggests that the Spanish learners tend to have an earlier alignment of rise-falls in their L2 American English. In their utterances the F0 reaches its peak very early (*before* the accented syllable), and falls just before and during the beginning of the accented syllable. This may have caused the American judges to conclude that the stress was placed incorrectly (too early), since Americans would expect the falling pitch to occur much later.

### 5.3.2. Pitch range and tonal scaling

Previous research has given some evidence that pitch range in L2 intonation is likely to differ from that of native speakers. Usually, the

pitch range in the intonation of L2 is reported to be too narrow (Backman, 1977; Jenner, 1976; Willems, 1982). But it is not so straightforward to interpret these findings, as different researchers used different methods to measure the pitch range. For example, Willems (1982: p.56-57) found that Dutch learners of L2 English, produced 'smaller excursions of pitch movements' on the rise, the non-final fall and the fall, expressed in semitones<sup>16</sup>. Backman (1977: p.32) on the other hand, measured the pitch range in Hertz, from the highest to the lowest point in the sentence. Such differences make it difficult to compare the findings in these studies. Another factor which makes comparisons difficult is the difference in the levels of proficiency of learners in these studies. Backman (1977) based her findings on the production of English by relative beginners, whereas the subjects in Willem's study are intermediate learners of English. However, based on these findings, it seems reasonable to assume that beginning and intermediate learners of an L2 may find adjusting the pitch range in the L2 difficult. The direction of that difficulty (i.e. whether learners will produce a bigger or smaller pitch range) is not clear, it probably depends on the extent to which pitch range in L1 and L2 is different.

Another difficulty when comparing pitch range across languages is that pitch shows a great deal of inter- and intra-speaker variation. So, speakers can differ a lot with respect to each other (e.g. individual differences in pitch range, or differences between female and male voices). For example, it has been found that two rises that are the same in size when measured in Hz are perceived as larger when produced by a lower pitched voiced than when produced by a higher pitched voice ('t Hart, Collier, and Cohen, 1990). On top of that, speakers can also modify their pitch range for communicative purposes, like expressing surprise, anger etc. As inter- and intra-speaker differences can obscure general patterns in intonation, it may be necessary to abstract away from such

---

<sup>16</sup> The term semitone is borrowed from music, and will be explained below.

differences, especially when comparing pitch range across languages or when studying L2 intonation.

Several researchers have tried to come up with a model for pitch range. To accommodate for perceptual differences like those mentioned above, pitch changes are sometimes expressed in the logarithmical semitone scale ('t Hart *et al.*, 1990) or in the psycho-acoustic 'Equivalence Rectangular Bandwidth' (ERB) scale (Hermes and Van Gestel, 1991; Hermes and Rump, 1994). Other models that seek to factor out sources of variation in pitch range are Earle's and Rose's normalising models (Earle, 1975; Rose, 1987). Earle specified each speaker's highest and lowest F0 values, by assigning them a value of 100 and 0 respectively. Then, every F0 value is expressed relative to these points, so in effect F0 is expressed on a percentage scale. In this way, Earle was able to characterise differences between lexical tones in Vietnamese words. A similar model is proposed by Rose (1987) in his study of Wu Chinese. However, instead of assigning a value to the top and bottom of each speaker's range, his normalisation model is based on z-scores.<sup>17</sup> In this way it is possible to discover regularities in inter-speaker variation of pitch range. Both Earle (1975) and Rose (1987) found a high degree of inter-speaker agreement in their normalised data.

The fundamental frequency level at which a H or L tone occurs is called scaling. Earle (1975) and Rose (1987) found that when expressed on a normalised scale, the different lexical tones in Vietnamese and Wu Chinese are highly systematically scaled. A number of studies found that the scaling of tones in European languages is also substantially invariant across speakers and across utterances. Regularities of this kind have been found in a number of studies (e.g. Bruce, 1982; Ladd, 1988; Ladd and Terken, 1995; Liberman and Pierrehumbert, 1984; Pierrehumbert and Beckman, 1988).

---

<sup>17</sup>Z-scores are used to express how many standard deviations a certain data point (in this case F0) is removed from that subject's mean.

It is possible that, just like alignment, the scaling of H or L tones is language-specific. And just as it may be difficult for language learners to acquire the specific alignment pattern of a foreign language, it may also prove difficult to acquire the characteristic scaling of the L2.

### 5.3.3. Word stress and nuclear accent placement

It is generally accepted that L2 learners often have difficulty with the correct placement of word stress, especially in the initial stages of the learning process. For example, Willems (1982) in his study of the intonation patterns of Dutch learners of English reports that

Although we had marked the desired positions of pitch accents in the written presentation of the dialogue, yet many native speakers of Dutch occasionally preferred displaced positions of pitch accents (Willems, 1982: p101).

Many other studies suggest that word stress placement is problematic for L2 learners (e.g. Adams and Munro, 1978; Archibald, 1992; Fokes and Bond, 1989; Wenk, 1985). Also, studies on the teaching of L2 prosody suggest (although based to a large extent on anecdotal evidence and impressionistic observations) that word stress needs to be given special attention in the classroom (e.g. Anderson-Hsieh, Johnson, and Koehler, 1992; Buysschaert, 1990).

Apart from difficulty with prominence within a word, L2 learners also seem to experience difficulty with the correct placement of prominence at the sentence level (e.g. Backman, 1979; Jenner, 1976). Just as a language can have phonemic contrasts, like a contrast between a voiced and a voiceless stop (/d/-/t/), the prominence system within a language is also a system of contrasts. A word is produced with more acoustic salience, or prominence, in order to contrast that word with other less prominent words. Just as phonemes serve to distinguish one word from another word, a system of prominence allows a speaker to contrast the relative importance of words.

Both Jenner (1976) and Backman (1979) report that language learners often move the most prominent word of the sentence (the main or nuclear accent) too far to the left in their L2 utterances. Again, it is not clear whether this is caused by a phonetic or a phonological error. Most of the test sentences Backman (1979) presents in her study consist of monosyllabic words only. If the Spanish learners of English have aligned the rise-fall in a sentence like "I'm late" too early, with the peak occurring just before the onset of the word "late", native Americans may have perceived this as a prominence on "I'm". This may have led to the perception of a shift of the nuclear accent to the left.<sup>18</sup> For this reason, these results have to be interpreted with caution.

Another reason for questioning the results obtained in the above mentioned studies, is the fact that the use of acoustic cues to signal stress may be different across languages. Beckman (1986), for example, suggested that even though languages use the same parameters to signal stress, their relative importance is language specific. For example, Americans use all four perceptual cues to stress (F0, duration, amplitude, and spectral coefficient) to the same extent, whereas Japanese use F0 cues to a much greater extent than other cues to stress (Beckman, 1986). As a consequence, when listening to American English, Japanese will rely mainly on F0 cues, and may disregard other cues to stress which should influence their perception of stress.

In production there also seem to be cross-linguistic differences in the cues used to signal stress. For example, Adams and Munro (1978) found a difference in the production of sentence stress between native and non-native speakers of English. Adams and Munro found that the "real difference between the stress production of the two groups lay not in the mechanisms they used to signal the feature [stress], but rather in their

---

<sup>18</sup> Unfortunately, it is not possible to inspect Jenner's (1976) data, as in his study no acoustic data are presented to support his conclusion.

distribution of it..." (p. 153). In a similar study Fokes and Bond (Fokes and Bond, 1989) found that much the same is true for word stress.

If it is true that the acoustic correlates of stress differ across languages, results of studies relying on native speakers' judgements of stress placement by non-native speakers have to be interpreted with caution. Native judges may presuppose certain acoustic cues to stress other than the ones produced by non-native speakers. It is therefore possible that the non-native speakers described in these studies do not actually produce errors in stress *placement*, but merely differ in the relative importance of the cues used to produce stress. A recent study by Low and Grabe (in press) seems to support this explanation. Their results indicate that the widely reported claim (based on native British English listener judgements) that British English and Singapore English differ in stress placement is not true. Their experimental data suggest that the apparent word-final stress in Singapore English (as opposed to the word-initial stress in British English) in words like *flawlessly*, is not the result of a difference of lexical stress placement. Instead, it seems that Singapore English and British English differ in the phonetic realisation of stress, with more phrase-final lengthening, and a lack of "deprominencing" in F0 in Singapore English than in British English. As a result, Low and Grabe argue that "the location of stress (or even its presence) cannot be judged impressionistically in any cross-linguistically valid way."

## 6. Aims and overview of the thesis

Following the discussion of the literature on the L2 acquisition of intonation, it appears that there are many issues that remain unresolved. First of all, the lack of research in this area is so striking, that it can only be fruitful to turn our attention to this aspect of L2 learning. Secondly, some of the few studies suffer from methodological problems. In this thesis, the

L2 acquisition of intonation will be further investigated, controlling for confounds seen in other studies.

It was thought that it would be interesting to examine L2 learners at the highest possible level of acquisition, as this may bring to light whether 'accent-free' (i.e. undistinguishable from native) production of intonation is achievable. The non-native speakers who participated in the experiments presented in this thesis had between 6 to 35 years of language experience.

The main aim of the thesis is to establish whether second language learners are able to acquire native patterns of L2 intonation, and whether this has an effect on the intonation pattern in their L1. The main hypothesis, tested in this thesis is that even after long exposure to the target language L2 learners will experience difficulties in the production of phonetic aspects of L2 intonation. It is further hypothesised that the better L2 learners become at approaching native L2 values, the more their L1 will be affected. For this reason, the subjects in the experimental studies are very advanced (near-native) speakers of the L2. The speakers are all native Dutch (L1) speakers of Modern Greek (L2).

A secondary aim of this thesis is to test whether a recent model, the Speech Learning Model developed by Flege and colleagues to explain *segmental* aspects of L2 learning, can also account for *suprasegmental* aspects of L2 learning, in this case intonation.

After this brief sketch of the main aims of the thesis, there now follows an overview of the remainder of the thesis.

In Chapter 2 a short overview will be given of the intonation in Dutch and Greek, and the differences between the two will be summarised.

In Chapter 3 three production experiments will be described, in which the peak alignment in prenuclear rising accents is examined. The first study (experiment 1) established cross-linguistic differences in peak alignment between Dutch and Greek. The second study (experiment 2) tested whether there were differences in the production of Greek prenuclear rising accents between non-native speakers and native

speakers of Greek. A further experiment (experiment 3) tested whether long exposure to a second language had an effect on the intonation of the L1. This experiment tested whether non-native speakers of Greek produced different peak alignment values in their Dutch than did Dutch speakers who had no extensive knowledge of Greek or any other foreign language.

In Chapter 4 two studies are reported which investigated the production of yes/no questions in the L2. The production of two types of Greek yes/no questions was investigated for non-native (Dutch) speakers of Greek. The first experiment (experiment 4) concentrated on Greek yes/no questions in which the main accent of the question is placed on the first content word of the sentence (nucleus non-final yes/no questions). The second experiment (experiment 5) investigated another type of Greek yes/no question and how it is produced by non-native (Dutch) speakers. In this type of yes/no questions the main accent is placed on the final word of the sentence. This experiment tested whether there were differences between native and non-native production of this type of Greek questions. Specifically, experiments 4 and 5 tested whether it was easier for L2 learners to acquire a 'new' intonation pattern, like that in experiment 4, or a 'similar' pattern, like that in experiment 5.

Chapter 5 presents the final conclusions of the thesis, as well as suggestions for further research.

## Chapter 2

### Intonation in Dutch and Greek

The aim of this chapter is not to present an exhaustive account of Dutch and Greek intonation, but rather to provide analyses of those intonation patterns which are the subject of this thesis. Since many descriptions of Dutch intonation are based on the approach developed at the Institute of Perception Research (IPO), this approach will be described first. Some of the IPO findings relevant to the present study will be presented, together with some opposing viewpoints from researchers adopting a different approach. In section 2 a description of Greek intonation is given. The chapter concludes with a comparison of Dutch and Greek intonation, and shows how the assumptions of the Speech Learning Model can be applied to the L2 acquisition of intonation.

#### 1. Dutch intonation

##### 1.1. The Grammar of Dutch Intonation

Before describing Dutch intonation, it is necessary to describe the approach to intonation which was developed at the Institute for Perception Research (IPO) in Eindhoven ('t Hart and Cohen, 1973; 't Hart and Collier, 1975; 't Hart, Collier, and Cohen, 1990; Cohen and 't Hart, 1967). The IPO researchers developed a Grammar of Dutch Intonation (GDI) which initially grew out of their concern with speech synthesis, but has developed since into a theoretical model of Dutch intonation. This model is characterised by a primary interest in the acoustic-phonetic aspects rather than the linguistic meaning or function of intonation. The model assumes that the intonation contours of a language are reflected in

a limited number of recurrent discrete pitch movements, which are actively controlled by the speaker of that language (Cohen and 't Hart, 1967). It is argued that speakers can distinguish between voluntary and involuntary changes in the periodicity of the vocal fold vibration, and that listeners are only sensitive to those F0 changes that have been intentionally produced by the speaker. The GDI claims that:

... when listening to an utterance [listeners] are not following its pitch period by period (...), but are only sensitive to a certain number of pitch events ('t Hart and Collier, 1975: p.238).

The first aim of the IPO researchers was therefore to find out which of the pitch fluctuations in speech are relevant to listeners. For that reason they developed a method in which involuntary fluctuations in the pitch contour are ironed out by means of analysis-resynthesis techniques. Simply put, the pitch contour of an utterance is replaced by an artificial contour, in which pitch movements are replaced by the smallest possible number of straight lines (so that small pitch fluctuations are removed), while still maintaining perceptual equivalence with the original contour. Perceptual equivalence with the original contour is possible, because even though the pitch contour of the original utterance can be manipulated, all other acoustic properties of the original utterance (spectral composition, temporal structure) remain virtually unchanged. This method is called 'close-copy stylisation', and an example is shown in Figure 2.1

Since the original contour and the 'stylised' contour are perceptually equivalent, all of the pitch characteristics of the latter contour must be considered perceptually relevant for the listener of that language.

Once the perceptually relevant pitch movements of Dutch (i.e. the straight lines) are established, they are classified on the basis of several criteria (direction, timing with regard to syllable boundaries, rate of change, and size), after which they are labeled. Table 2.1 lists the relevant pitch movements with their labels.

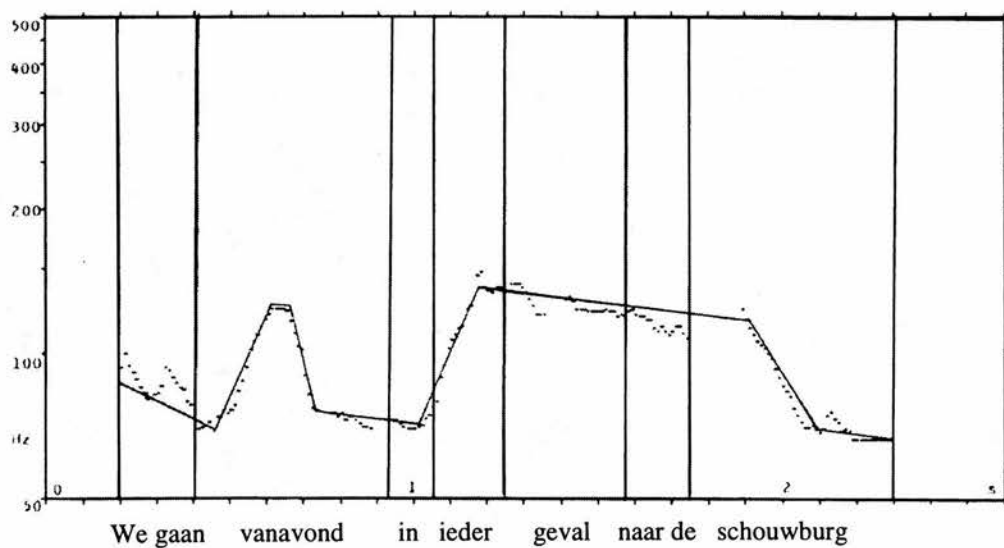
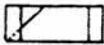
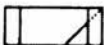

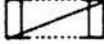


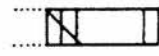

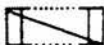
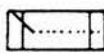


Figure 2.1. An example of a fundamental frequency curve (dotted line) with the stylised contour (solid line) superimposed (adapted from Willems, 1982: p. 40).

Table 2.1. *Description of the perceptually relevant pitch movements of Dutch according to the GDI (from Collier & 't Hart, 1981).*

	1	accent-marking rise, steep and early
	2	non-accent-marking rise, steep and very late
	3	accent-marking rise, steep and in the middle of the syllable
	4	gradual rise, marking more than one syllable
	5	short, extra rise, sometimes after 4
	A	accent-marking fall, steep and in the middle of the syllable
	B	non-accent-marking fall, at syntactic boundary, steep, between syllables
	C	non-accent-marking fall, steep and very late in the syllable
	D	gradual fall, usually not distinct from A in l&A
	E	half-fall, weakly accent-marking if occurring in isolation, steep and early in the syllable. It is sometimes displaced rightwards.

The close-copy in Figure 2.1 illustrates two variations of the 'hat pattern', a 'pointed hat' followed by a 'flat hat'. The 'flat hat' (which is depicted in the second half of the contour) is one of the best known basic intonation patterns of Dutch. It is represented as '1A', that is a 'type 1 Rise' (an early 'prominence-lending' rise), followed by a 'type A Fall' (a prominence-lending steep fall in the middle of the syllable), with a level high stretch ( $\emptyset$ ) in between.<sup>1</sup>

From this basic hat pattern other forms can be derived, for example the 'pointed hat' in which the rise and fall both occur on a single syllable. This pattern is represented as '1&A', an example of which is shown in the first half of Figure 2.1. Altogether, the GDI found that five different rises and five falls appear to be sufficient to describe Dutch intonation. A grammar of Dutch intonation has been formed which describes how these various pitch movements can be combined to form entire contours ('t Hart, Collier, and Cohen, 1990; Cohen and 't Hart, 1967).

In Table 2.1 it can be seen that rises and falls can be of two types, prominence-lending and non-prominence-lending. According to the GDI a pitch movement that occurs on a lexically stressed syllable lends perceptual prominence to the syllable in which it is realised. Each lexically stressed syllable of a prominent word will receive such a pitch accent. In contrast, there are pitch movements which do not lend prominence to a word or syllable. These non-prominence-lending pitch movements either extend over several syllables, or occur at phrase boundaries. This distinction between prominence-lending and non-prominence-lending pitch movements is quite similar to other approaches which describe intonation within autosegmental-metrical terms. In an autosegmental-metrical approach an accent-lending pitch movement is analysed as a

---

<sup>1</sup>Although in the IPO tradition stretches of level pitch are also labeled, by using 0 for low level stretches and  $\emptyset$  for high level stretches, some researchers (e.g. Caspers, 1994; Gussenhoven, 1984, 1988) have dispensed with this notation when describing basic intonation patterns in Dutch. I will omit notation of the level stretches, except when it is

*pitch accent*, which is associated with a stressed syllable. The non-accent-lending pitch movement that occurs at phrase boundaries is categorised as a *boundary tone*.<sup>2</sup> It is associated phonologically with some kind of boundary, such as the end of an utterance. Boundary tones can be either high (H) or low (L), and are often indicated as H% or L% (e.g. Pierrehumbert, 1980).

### 1.1.1. Criteria for classification

Apart from distinguishing rises and falls, the GDI also classifies pitch movements on the basis of other criteria, i.e. timing, rate of change, and size. These categories are not just labels, but explicitly specify how the movements are realised. For example, the 'fast, early, full rise' (Rise '1') has a duration of 120 milliseconds (ms) and an excursion size of 6 semitones (ST), its peak is reached 50 ms after the vowel onset ('t Hart, Collier, and Cohen, 1990). These acoustic definitions of each pitch movement are derived from further stylisation of the close-copies, aimed at replacing the different pitch movements with standard specifications for their various parameters. The standardisation process enabled the IPO researchers to make generalisations and group the perceptually relevant movements into a restricted number of categories. Furthermore,

---

essential for clarity.

<sup>2</sup> Another type of non-prominence-lending pitch movements which is recognised by the GDI is a pitch movement which does not occur at a phrase boundary, but which spreads over several syllable, like the 'type 4 Rise' and 'type D fall'. These pitch movements are not prominence-lending, but they are not boundary tones either. In some research based on the autosegmental-metrical approach to the study of intonation they are analysed as interpolations from a Low tone to a High tone, or vice versa (Cruttenden, 1992; Ladd, 1983).



standardisation provided the IPO researchers with a set of definitions for use in speech synthesis.<sup>3</sup>

In the next sections, the criteria employed by the GDI for the classification of pitch movements are discussed separately.

#### 1.1.1.1. Rate of change

The GDI distinguishes between *abrupt* and *gradual* pitch changes. It is thought that differences in the slope of pitch movements are only perceived by listeners when the duration of the movement is at least 250 ms. In other words, a difference in slope is only perceived when the pitch changes gradually over several syllables. Changes within a syllable are not thought to be noticeable ('t Hart, 1976).

However, in their description of relevant pitch movements in Dutch (Table 2.1), Collier & 't Hart (1981) note that the difference between the steep 'A' and the gradual 'D' is not thought to be distinctive in '1&A', even though the former takes place within a syllable and the latter extends over several syllables. Collier and 't Hart (1981) do not offer an explanation for this fact.

An explanation for the lack of perceptual difference between the two types of fall is given by Gussenhoven (1988), who proposes a different approach to describing intonation than the GDI. Gussenhoven's model is based on the autosegmental-metrical approach to intonation. The GDI treats pitch *movements* as the basic descriptive units of intonation,

---

<sup>3</sup> 't Hart and Collier (1975) actually recognise that their acoustic definitions are values they have settled on for speech synthesis purposes, and that they may not be accurate. For example, they state that "... some pitch movements occurred earlier or later in the syllable than is specified in their definition...[but they] have not attempted to incorporate such possible systematic discrepancies into a new version of the grammar" ('t Hart & Collier, 1975: p. 247-48).

whereas the autosegmental-metrical approach treats high (H) and low (L) *tones* as the fundamental units.<sup>4</sup>

Gussenhoven (1988) showed that the variety of Dutch intonation contours can effectively be described in terms of strings of H and L tones. He proposes three 'tonal morphemes' to account for all intonation contours in Dutch. The three morphemes that he proposes for Dutch are H\*L, H\*LH, and L\*H. Each accented position in a sentence is associated with one of these tonal morphemes. The asterisk (\*) indicates the central or accented tone, and it is associated with the accented syllable. The other elements of the tonal morphemes are associated with following unstressed syllables. The three morphemes can be modified by several rules and modifications, one of which - the tone linking rule - can explain the resemblance between the 'A' and 'D' in '1&A'. According to Gussenhoven, in a sequence of two H\*L morphemes (or in GDI terminology a sequence of two pointed hats, '1&A...1&A'), partial linking of H\*L...H\*L will change the steep fall (Fall 'A') into a gradual fall (Fall 'D'), as is illustrated in Figure 2.2. He argues that because partial linking is not very different from no linking, it is not surprising that these two variants of H\*L...H\*L are perceptually similar.

---

<sup>4</sup>The autosegmental-metrical theory is generally associated with the work of Pierrehumbert (e.g. Pierrehumbert, 1980; Pierrehumbert and Beckman, 1988), although earlier work has influenced the development of the theory (Bruce, 1977; Goldsmith, 1976; Leben, 1973; Leben, 1976; Liberman, 1975). For a description of this approach, the reader is referred to Ladd (1996).

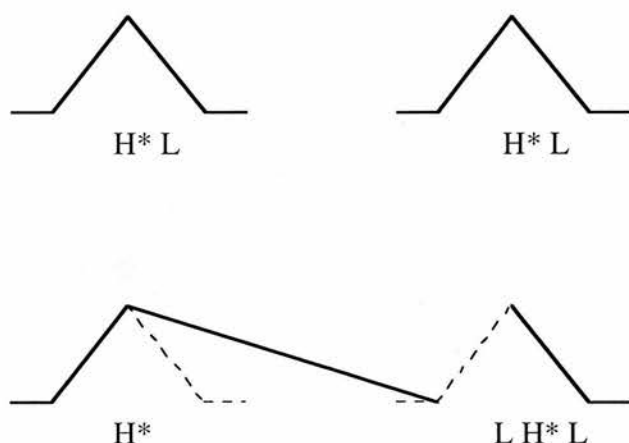


Figure 2.2: Illustration of two 'pointed hats' ( $H^*L...H^*L$ ), which are unlinked in (a), and partially linked in (b). The partial linking in (b) turns the steep fall (type 'A') into a gradual fall (type 'D').

#### 1.1.1.2. Size of pitch movement

A second criterion which the IPO researchers used in the standardisation of pitch movements is the *size* of the pitch movement. The GDI distinguishes between full and half sizes of movements. This distinction is based on 't Hart (1981) who attempted to determine the sensitivity of the human ear to differences in the size of pitch movements. He found that differences of at least 3 semitones are necessary to distinguish on the basis of excursion size. Other researchers provided evidence that there is a significant correlation between the size of pitch movements and the perception of prominence (Gussenhoven and Blom, 1978; Van Katwijk, 1974). However, the GDI distinguishes only one half-size movement, all other movements are considered to be full sized. The only half-size movement recognised by the GDI is the half-fall 'E'. It is called a half-fall since it makes the impression of being neither high nor low ('t Hart and Cohen, 1973). It is further defined as weakly accent marking, steep and occurring early in the syllable, although it can be displaced to the right (see also table 2.1). As it is thought that the half-

fall 'E' does not usually occur in the intonation patterns which are the subject of this thesis (prenuclear accents and yes/no questions), the size of pitch movements will not be discussed any further here. For a different analysis of the pitch movement 'E' the reader is referred to Gussenhoven (1988).

### 1.1.1.3. Timing

As we have seen in Chapter 1, section 5.3.1, in the last decade or so, phonologists have become more aware of the importance of pitch alignment. The GDI had already recognised the relevance of alignment differences for a variety of pitch movements long before this interest in alignment started to grow. The GDI distinguishes between three types of pitch movements in Dutch: *early*, *late*, and *very late*. Support for this three-way distinction was found in an experiment by Collier (1975). His results, which will be further discussed in section 1.2, suggest that listeners classified contours into three groups corresponding to the location of the pitch movement within the experimental syllable.

Other perception experiments showed that alignment plays a role in the perception of prominence. For example, Govaert and Van Katwijk (1968) systematically varied the position of a pitch rise (type '1') and fall (type 'A') in synthesised nonsense utterances. Listeners had to indicate which syllable they perceived as prominent. The results indicate that for a rise to be perceived as prominent it should be situated rather early in the syllable, a fall however is perceived as more prominent when it is located rather late in the syllable. A similar experiment was reported in Collier (1972). In this experiment subjects were asked to make a target syllable of different stimulus sentences as prominent as possible, by adjusting the location of the pitch rises. His results also suggest that a syllable is perceived as most prominent when the rise is positioned early.

### 1.1.2. Pitch contours

According to 't Hart *et al.* (1990) the individual pitch movements (rises and falls) that are recognised in the GDI combine into pitch contours. The internal structure of these contours is expressed in a grammar. This grammar defines which sequences of pitch movements are permitted. According to this grammar (the GDI) individual pitch contours are first grouped into pitch configurations (like 1A, 2B, etc.), which are then grouped into pitch contours. According to the GDI, pitch contours tend to coincide with clauses or complete utterances ('t Hart *et al.*, 1990:p.151). The proposed grammar was verified in a corpus of speech and it appeared that only 6% of the contours were unaccounted for. As further details about this grammar are not relevant for this thesis, the interested reader is referred to 't Hart *et al.* (1990).

### 1.2. Alignment in Dutch<sup>5</sup>

In Chapter 1 it was seen that alignment (i.e. the location of H or L tones with respect to the segmental string) is a possible candidate for errors in L2 intonation. It is therefore important to describe alignment in native Dutch. The alignment of pitch movement in Dutch was investigated in a perception experiment by Collier (1975). He presented listeners with synthetic intonation contours in which the rise was shifted through the utterance's final syllable (in steps of 20 ms) and asked them to decide which utterances had similar pitch contours. It was found that listeners classified the pitch contour into three categories depending on the location of the pitch movement. In this study the precise location of the rise is not specified. However, in other studies these timing specifications are expressed as the location of the endpoint of the rise with

---

<sup>5</sup> This section owes much to Verhoeven's (1991) discussion of the literature on alignment in Dutch. As suggested in Chapter 1, it is thought that alignment may be a possible source of

respect to the vowel onset of the syllable. The endpoint of the early rise (rise '1') is preferably situated 30 ms on average after the vowel onset, the late rise (rise '3') at 90 ms after the vowel onset, and the very late rise (rise '2') is said to come as late as physically possible ('t Hart and Collier, 1975).<sup>6</sup>

Differences in the position of the rise within a syllable were also subject of investigation in Boves, ten Have, and Vieregge (1984). Boves *et al.* were interested in the timing differences between the prominence-lending pitch rises '1' and '3', of which the former is said to occur early in the syllable, whereas the latter occurs late. In their experiment subjects were asked to imitate sentences containing clear examples of these early and late movements. The contours, which were produced by a highly trained specialist in GDI, were instances of both the hat ('1A') and the cap ('3C') pattern. In Figure 2.3. an example of the hat and the cap pattern is shown.

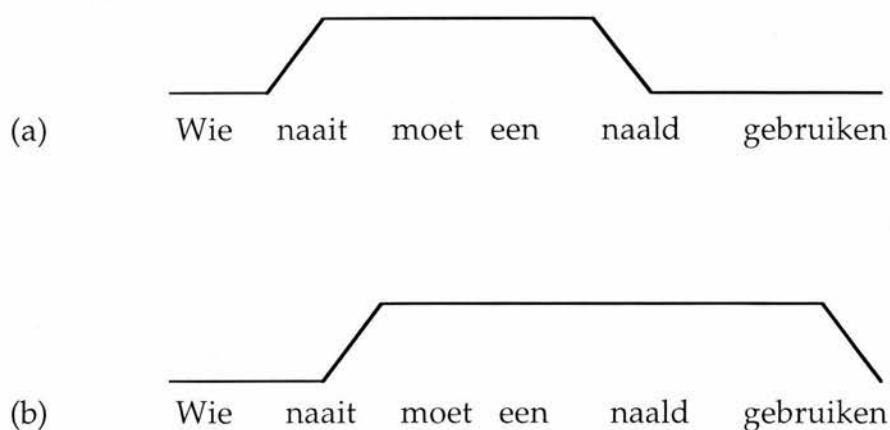


Figure 2.3: Illustration of the hat (a) and cap (b) pattern in the test items used by Boves *et al.* (1984).

---

error in L2 intonation. Therefore, the findings on alignment are elaborated in detail, whereas other criteria for classification are only summarised.

<sup>6</sup> As mentioned earlier, these timing specifications were mainly intended for speech synthesis purposes and 't Hart and Collier (1975) actually recognise that they may not be accurate (see also footnote 21).

In the hat pattern (panel a) the prominence-lending rise '1' occurs early in the syllable, whereas in the cap pattern (panel b) the prominence-lending rise '3' occurs late. The difference between the onsets of the two types of rise in the test items is approximately 60 ms. Apart from a difference in the type of rise, the contours also differ in the location and type of fall in the second part of the utterance. In (a) the fall is of type 'A', occurs on the fifth syllable, and is prominence-lending. In (b) the fall is of type 'C', it is non-prominence-lending, and occurs at the very end of the utterance. Two subjects were presented with stimulus tapes with the test sentences, and were asked to imitate these sentences on tape. From these imitations F0 and alignment measurements were taken. The alignment was expressed as the onset of the rise with respect to the vowel onset. From the results it appears that subjects were unable to imitate the alignment differences which were present in the test items. Instead, the alignment differences were replaced by small differences in excursion size and slope.

As a follow-up of their production experiment, Boves *et al.* (1984) also designed a perception experiment. In this experiment three naive and three highly trained subjects were presented with the imitations of the two test utterances, and asked to label the type of rise. The results indicate that the two groups of subjects show extreme differences in performance. The trained subjects were able to distinguish the different types of rise with 100% accuracy, whereas the naive subjects did not perform above chance. The authors conclude from these results:

... all imitations constituted acceptable realizations of hat or cap patterns, as all three 'trained' judges reported that they had based their decisions entirely on the contour as a whole (Boves, Ten Have, and Vieregge, 1984: p.33)

When the subjects were prevented from taking recourse to the differences in the second halves of the contours, by presenting them with stimulus tapes which only contained the first two syllables of the imitations, the trained listeners performed less well. It thus seems that familiarity with GDI can strongly influence judgements. The trained

judges are forced by the GDI to identify the rise as type '1' or '3' on the basis of the final part of the contours, in which the differences are easily noticeable. When the final part of the contour is absent, differences between the judgements of highly trained and untrained listeners start to disappear.

In most of the studies described above, the timing specifications are expressed as the location of the endpoint of the rise. For example, Collier (1970, as reported in Collier 1972) conducted an experiment in which subjects had to adjust the location of the pitch rise in several stimulus sentences so as to make the target syllable as prominent as possible. It was thought that the position of the pitch rise could be influenced by the phonetic structure of the prominent syllable. Therefore, its phonetic structure was varied (with an onset of zero to three consonants, followed by the vowel /a:/ or /a/). According to Collier (1972) the results of this experiment indicate that

... the phonematic structure of the prominent syllable (viz. zero to three consonants preceding the vowel /a/ or /a/) does not significantly influence the location of the pitch rises. The only relation of the pitch rise to the phonematic structure apparently is that subjects tend to relate its position to the vowel onset (CV transition) and not to the beginning of the syllable (Collier, 1972: p.83).

Collier (1970, as reported in Collier 1972) also found that when adjusting the location of pitch rises, listeners paid more attention to the position of their peak (with respect to the vowel onset) than to their onset. In other words, in order to decide whether a syllable is prominent or not, the position of the peak is more important than the onset time of the rise. This was also confirmed by a pilot experiment (which Collier does not describe in detail), where the onset moments of pitch rises of different lengths revealed a wider spread (280 ms) than the positions of the peaks (90 ms). This suggests either that the peak of pitch rises in Dutch is more consistently aligned than is the onset of the pitch rise, or that listeners are more sensitive to the alignment of peaks rather than valleys. If the former is true, it suggests (even though the GDI does not say so

explicitly) that the peak in a type '1' rise is actually 'anchored' to the stressed vowel.

This specific suggestion was challenged by Caspers (1994; 1993). She tested the influence of time pressure on the shape, pitch level and segmental alignment of the pitch rise '1' and the fall 'A'. Three types of time pressure were used: normal versus fast speech, long versus short target vowels, and absence versus presence of competing pitch movements. Caspers expected that the rise would be rather invariantly aligned, with its offset anchored relative to the vowel onset (as suggested by the GDI). However, no such anchoring was found. There was no evidence for a fixed relationship between the offset of the rise and the vowel onset. Instead, she found that the position of the onset of the rise was relatively insensitive to time pressure. She concluded that it is not the offset but rather the onset of the rise which is anchored in the segmental structure.

Apart from this production study, Caspers (1994) also reports a perception experiment in which resynthesized pairs of utterances containing a 'low-anchored' or a 'high-anchored' pitch rise were presented to listeners. In a 'low-anchored' rise the onset of the rise is fixed and coincides with the syllable onset; the 'high-anchored' rise has its offset at a fixed position of 50 ms after the vowel onset. Alignment differences were created by varying the duration of the rise and the initial consonant. The target syllables used in the experiment were /mən/, /ma:n/, /pən/ and /pa:n/, and the contour type which was used was a flat hat. Twenty-five naive and twenty five experienced listeners were presented with the different stimulus pairs, and were asked to indicate whether the members of the pair were the same or different. After this discrimination task the listeners were asked to indicate for each of the stimulus pairs which item they preferred ('forced choice preference test'). The results indicate that listeners are able to perceive differences between a low and a high anchor, and that they prefer a low over a high anchor in a type '1' rise (in 65% of the cases).

### 1.3. Shortcomings of alignment studies

The distinction the GDI makes between three locations of alignment (early, late, and very late) is not without problems. Most of the studies described in the previous section do not take into account that there are other factors which may influence alignment. In Chapter 1, it was seen that a variety of prosodic conditions can influence the position of the peak, like speech rate, the proximity of a following word boundary, and the number of unaccented syllables intervening between the test syllable and the following accent (e.g. Silverman and Pierrehumbert, 1990; Prieto *et al.*, 1995, Arvaniti *et al.*, 1998). For example, Silverman and Pierrehumbert (1990) found clear parallels between the effects of prosodic context on the alignment in prenuclear and nuclear H\* pitch accents. However, there remains a difference in the timing of the peak in prenuclear and nuclear accents in English, i.e. when measured in absolute terms the peak in nuclear position occurs earlier than in prenuclear position (Silverman and Pierrehumbert, 1990: p. 96). The alignment differences between early, late, and very late locations found by Collier (1975) was based on an experiment with nuclear accents only. The fact that listeners distinguish between early, late and very late position of the rise in a nuclear accent, does not necessarily mean that listeners also perceive these differences in other prosodic contexts. In fact, Verhoeven (1991, 1994) found that there is no justification for the conclusion that listeners are able to perceive this kind of alignment differences in prenuclear accents. He showed that there is no indication of categorical perception of rise alignment in prenuclear accents: listeners were unable to distinguish between an early (type '1') and late (type '3') rise.

Furthermore, the studies on alignment differences in *nuclear* accents, present rather inconsistent results. On the one hand, the GDI suggest that the *offset* of a type '1' rise is invariantly aligned with respect to the segmental structure (e.g. Collier, 1970b; 't Hart *et al.*, 1990). On the other hand, Caspers' (1994) results seem to suggest that it is the rise *onset* which

is anchored in the segmental structure. For the time being, this issue remains unresolved.

#### 1.4. Prenuclear accents in Dutch

Most of the studies described in the previous sections have studied the perception of Dutch nuclear accents. The term nuclear accent is taken here to refer to the main pitch accent of an intonational phrase, i.e. the one that is most prominent. It has elsewhere been referred to as 'nucleus' (Palmer, 1992), 'tonic' (Halliday, 1967), and 'primary stress' (Cruttenden, 1968), to name just a few. The nuclear accent is usually thought to be the last pitch accent in an intonational phrase. Pitch accents which are non-final in an intonational phrase, are called prenuclear accents.

Languages differ in their preference for the type of prenuclear accents. For example, Dutch shows a preference for rising pitch accents, whereas Welsh prefers low prenuclear accents (Cruttenden, 1968).

The Dutch rising prenuclear accents are usually of the type '1&D', that is a steep, accent-lending rise, followed by a gradual fall, as shown in Figure 2.4. Gussenhoven (1988), as we saw above, analyses it as a H\*L accent, which is partially linked with the following H\*L accent (which was shown in Figure 2.2). As was pointed out in section 1.1.1.1, the gradual fall 'D' is usually not distinct from the steep fall 'A' in '1&A'. It is thus likely that differences between a '1&D' and a '1&A' are not perceived by listeners.

On the basis of the available literature, it is not clear how the rise is aligned in prenuclear accents of the type '1&D'. Collier (1975) distinguishes between early, late and very late alignment, but his conclusion is based on the perception of pitch rises in an utterance-final syllable. Verhoeven (1991; 1994) did not find any evidence for a categorical difference between an early and late rise in prenuclear

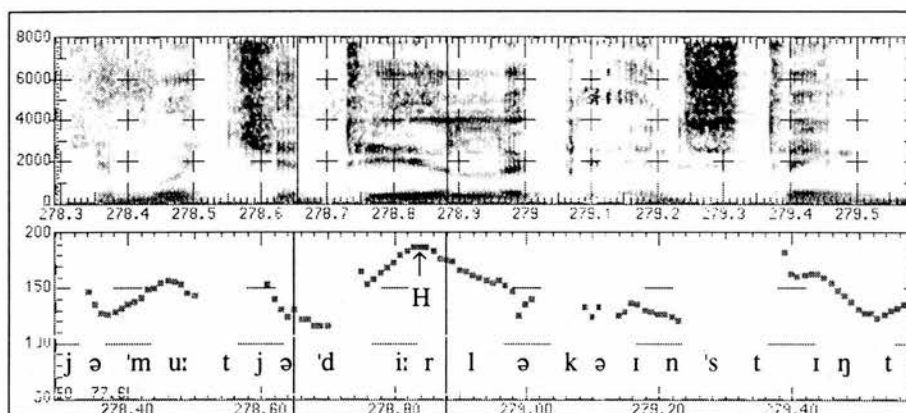


Figure 2.4. Spectrogram and F0 trace of the Dutch pre-nuclear accent 'Je moet je dierlijke instinct [niet altijd onderdrukken]' "You should not always suppress your animal instincts". The vertical lines delimit the stressed syllable of the word bearing the pre-nuclear accent.

accents. However, his findings were based on perception experiments in a 'flat hat' (type '1ØA'). It is unclear how his findings relate to the alignment of the rise in a pointed hat or '1&D'. Therefore, on the basis of the literature available, no predictions can be made as to the position of the rise in prenuclear '1&D' accents.

### 1.5. Yes/no question intonation in Dutch

In Dutch, we may distinguish three types of questions: Wh-questions, yes/no questions, and declarative questions. In wh-questions, interrogative words (like *who*, *when*, *where*, etc.) are used to mark interrogativity. Yes/no questions are usually marked by syntactic means, specifically by inversion of subject and finite verb. Declarative questions are not marked by either syntactic or lexical means, and are distinguished by intonation alone from otherwise identical statements. It has been shown by Haan, Van Heuven, Pacilly, and Van Bezooijen (1997) that the intonation pattern in yes/no questions is fairly similar to that in declarative questions. Therefore, in this thesis (when referring to Dutch question intonation), the term yes/no questions will be used to refer to both yes/no and declarative questions (unless stated otherwise).

Dutch yes/no question intonation is often described as hammock-shaped, with a high beginning, an equally high ending, and a low stretch in between (e.g. Daan, 1938; Es, 1932, as cited in Haan *et al.* 1997). The most obvious feature of Dutch yes/no question intonation is its rise at the end of the utterance. In the GDI terminology it is called the non-accent-lending rise '2'. In autosegmental-metrical terms it is referred to as a high boundary tone (H%).

The GDI does not distinguish between, for example, statements and questions. This is a deliberate choice of the IPO researchers, whose approach to intonation is bottom up, rather than top down, and meaning or function do not play a role in their analysis of Dutch intonation (e.g. 't

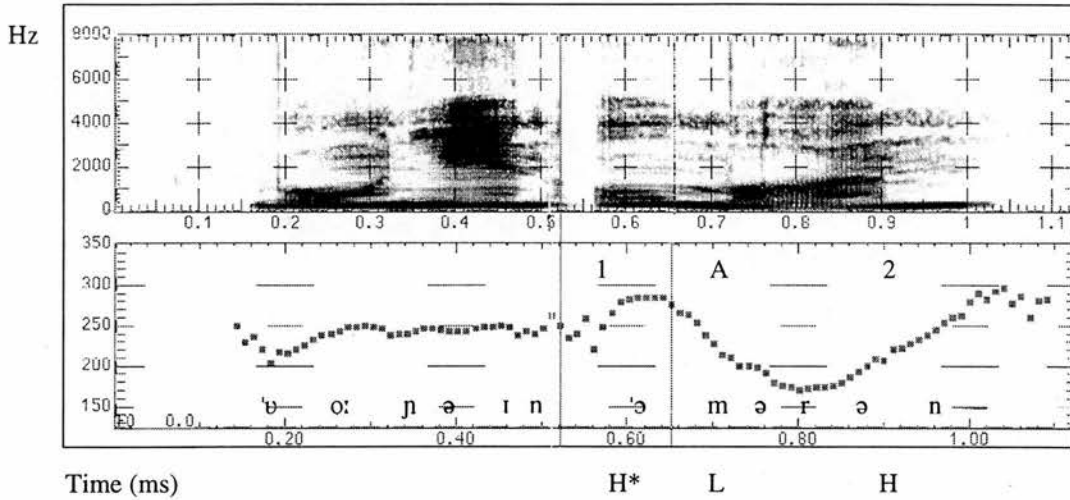
Hart & Collier, 1975; 't Hart *et al.*, 1990). For this reason, no explicit reference to Dutch yes/no question intonation can be found in the GDI. However, if one were to analyse Dutch yes/no question intonation in terms of the GDI, Dutch yes/no question intonation would have to be analysed as an accent-lending rise, followed by an accent-lending fall, and a final rise at the end of the utterance ('1A2'), or as an accent-lending fall followed by a final rise ('A2') (Collier & 't Hart, 1981).

In other analyses of Dutch intonation meaning or functional aspects of intonation have been taken into account (e.g. Caspers, 1998a; Caspers, 1998b; Gussenhoven, 1984; Haan, 1999; Haan *et al.*, 1997). For example, Gussenhoven (1988) provides an analysis of Dutch yes/no question based on the autosegmental-metrical approach to intonation. He analyses what the GDI would call the '1A2' as the sequence H\*LH; the 'A2' is analysed as the sequence L\*H. In both cases the starred tone is associated with the nuclear or main pitch accent of the phrase. In other words, the nuclear accent in Dutch yes/no questions can either be high (H\*) or low (L\*). If it is high, F0 is high on the stressed syllable of the most prominent word in the utterance, and begins to drop in that same syllable (Collier & 't Hart, 1981).<sup>7</sup> If it is low, the F0 drop starts earlier, before the accented syllable (Gussenhoven, 1988). In both cases it remains low until late in the final syllable where it usually starts rising again ('t Hart & Cohen, 1973). An example of these two yes/no question contours is given in Figure 2.5, panels (a) and (b). The GDI analysis is given above the contour, the autosegmental-metrical analysis below it.

---

<sup>7</sup> However, it is also possible that F0 does not drop after the H\* accent. In that case, the H% boundary tone immediately follows the H\* accent. According to Haan (personal communication) this H\*H% sequence is relatively frequent, and was observed in approximately 25% of her corpus of yes/no and declarative sentences. In terms of the GDI this sequence would have to be analysed as a '1 2', a combination of two rises which in fact is not allowed in their grammar (Gussenhoven and Rietveld, 1997).

(a)



(b)

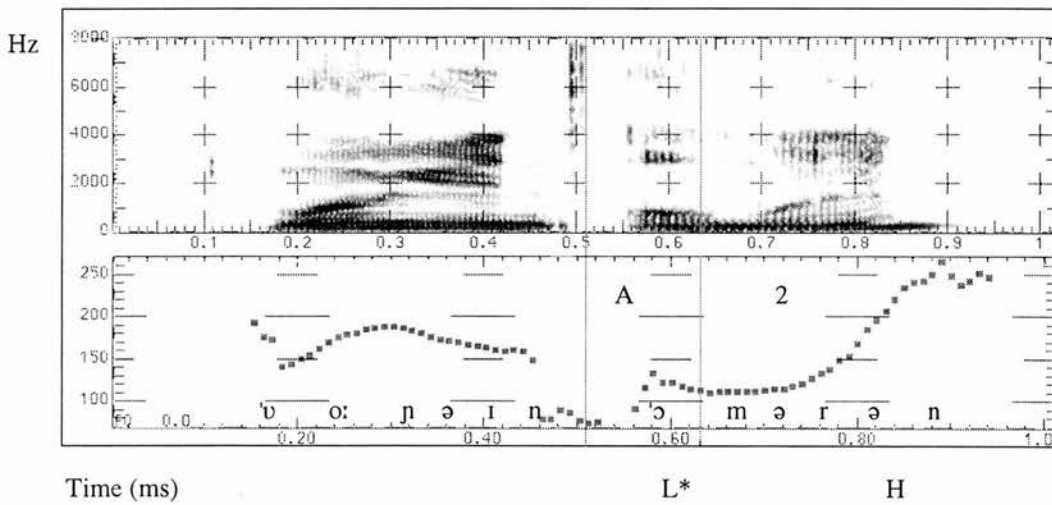


Figure 2.5. Spectrogram and F0 trace of the Dutch yes/no question 'Woon je in Ommeren?' 'Do you live in Ommeren?' Panel (a) shows an example of a H\*LH (or '1A2') contour. Panel (b) shows an example of a L\*H (or 'A2') contour. The vertical lines delimit the stressed syllable of the nuclear word.

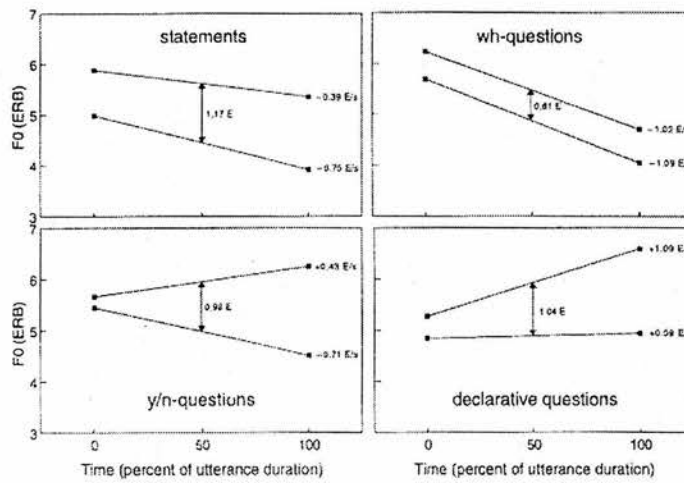


Figure 2.6. Slope of upper and lower regression lines (in ERB per second) for statements, wh-questions, yes/no questions, and declarative questions (adapted from Haan et al., 1997).

Characteristics of Dutch question intonation have also been investigated by Haan, Van Heuven, Pacilly & Bezooijen (1997). They found that apart from a steep final rise, which is present in the majority of Dutch questions, another characteristic of Dutch question intonation is its higher and narrower register when compared to statements. Figure 2.6. shows the global F0 differences between Dutch statements, wh-questions, yes/no questions and declarative questions.

## 2. Greek intonation

Greek intonation has been much less thoroughly investigated than Dutch. Due to the lack of technical facilities at the time they were written, most of the studies on Greek intonation are based largely on auditory observations by the authors, and not on experimental analysis (e.g. Setatos, 1974; Tsakirides, 1980; Waring, 1976; Waring, 1982). Unfortunately, a correct auditory analysis of an intonation contour is not an easy task, even for an experienced listener (especially if the speaker is a non-native speaker of Greek). This has led to the rather sketchy descriptions of Greek intonation in some textbooks on Greek (e.g. Holton, Mackridge, and Philippaki-Warburton, 1997; Joseph and Philippaki-Warburton, 1987; Mackridge, 1985).

The most elaborate of these auditory studies on Greek intonation was produced by Waring (1976, 1982). Waring recognises three basic pitch levels (high, mid, and low) and five basic tones (fall, rise, rise-fall, and fall-rise). In addition to these basic tones he specified a special 'raised-fall'. This final fall to mid pitch was thought to be almost exclusively restricted to yes/no questions. However, Greek yes/no questions are not always realised with this intonation contour. In fact, Waring (1976, 1982) points out that to Greeks the clearest way of indicating a yes/no question is the use of a 'rise-fall'. But, even though Waring recognised that the most common pattern for Greek yes/no questions is the 'rise-fall', it was the

'raised fall' which was interpreted as the most important characteristic of yes/no questions in most of the textbooks on Greek (Holton, Mackridge, and Philippaki-Warburton, 1997; Joseph and Philippaki-Warburton, 1987; Mackridge, 1985). However, such a 'fall to mid pitch' was not observed in a recent experimental study on Greek yes/no question intonation by Arvaniti, Ladd and Mennen (forthcoming). Instead, they report that yes/no questions are usually associated with a rise-fall, which falls as low as the final low in statements.

Botinis (1989b) in his work on the phonetic correlates of stress in Greek, is the first major published study that provides us with some instrumental data on Greek intonation. His work is to a great extent influenced by Bruce and Gårding's analysis of Swedish prosody (cf. Bruce and Gårding, 1978). With regards to intonation, Botinis found that one of the correlates of stress is a rising fundamental frequency. When stress is in prefocal or focal position, this rise starts at the consonant onset of the lexically stressed syllable. The highest point of the rise occurs at the end of the stressed syllable or in the first post-stressed syllable, when stress appears in prefocal position. However, when the stress appears in focal position, the peak occurs in the second half of the stressed syllable. No rise is observed in postfocal position, where fundamental frequency seems to flatten. It appears that, when stress occurs in postfocal position, fundamental frequency is not a cue to word stress.

Botinis (1989a) also investigated some aspects of Greek discourse intonation. He recognises several turn-keeping cues in Greek intonation. Speakers can convey the message that they intend to keep their turn, by producing a large pitch rise during the final part of the unit. Such a pitch rise can occur on stressed as well as unstressed syllables, and is sometimes called a 'continuation rise'. These continuation rises seem to be larger than the pitch rises associated with focal elements (i.e. speech elements which stand out as more 'informative' from the rest of the unit are said to be in focus) (1989a). On the other hand, when speakers intend to close the topic or conversation, a low pitch or a pitch fall was observed. The

association of low or falling pitch with completion, and the association of high or rising pitch with questions and non-finality have been reported for many other languages, and have led many investigators to believe that there is a universal common core for (some aspects of) intonation (e.g. Bolinger, 1978; Brazil, 1985; Brown, Currie, and Kenworthy, 1980; Brown and Yule, 1983).

Recently, a study on the meanings and functions of Greek yes/no questions has appeared, which is partly based on instrumental data. Papazachariou (1997) investigated Greek yes/no question intonation from a sociolinguistic point of view. More specifically, he was concerned with the sociolinguistic role of intonation among adolescents in Northern Greece. He defined six meaningful intonation variables that appear on one word yes/no questions, together with their function in conversation. Furthermore, the origins of the variants of each variable are studied, as well as their correlation and interaction with regional and demographic parameters, and with the social construction of gender identity. The six meaningful intonation variables Papazachariou (1997) recognises in yes/no questions, are 'high-falling', 'low-falling', 'mid-falling', 'high-rising', 'low-rising', and 'mid-rising'. The definition of these intonational units is in part based on previous work on the pragmatic function of intonation, specifically that of Gussenhoven (1984). It appears that the high-falling and high-rising variables indicate some form of uncertainty on the part of the speaker about the truth-value of the utterance. Low-falling and low-rising variables, on the other hand, indicate that the speaker is certain about the truth of the propositional content of the utterance. Mid-falling and mid-rising patterns indicate that the speaker does not have a subjective opinion about the truth-value of the utterance.

## 2.1. Prenuclear accents in Greek

Non-final or prenuclear accents in Greek declarative intonation are characterised by a rise that begins at the onset of the accented syllable and reaches its peak on the *following* syllable, during which F0 starts falling again (Arvaniti, Ladd, and Mennen, 1998; Botinis, 1989b; Dauer, 1980). Figure 2.7 shows an example of such a contour. If one compares this contour with that in Dutch prenuclear accents shown in Figure 2.4 (above), it can be seen that there is some phonetic resemblance between the two. However, there seems to be a difference in the timing of the peak. In Dutch prenuclear accents the peak seems to occur *within* the accented syllable, whereas in Greek it consistently occurs in the *following* vowel. Specifically, Arvaniti *et al.* (1998) found that in Greek prenuclear accents with test words with antepenultimate stress, the peak (H) occurs around 10 to 20 ms after the beginning of the first postaccidental vowel.

In fact, Arvaniti *et al.* (1998) found that it is not only the H which is invariantly anchored to a specific point in the segmental string. In this and a previous study, they found that the L target in Greek prenuclear rising accents is also fixed relative to the segmental string; specifically, the L is aligned just before the onset of the first consonant of the accented syllable (Arvaniti and Ladd, 1995; Arvaniti, Ladd, and Mennen, 1998).

Arvaniti *et al.* (1998) also found evidence for prosodic effects on alignment, like those found by Silverman and Pierrehumbert (1990) and Prieto *et al.* (1995). In Arvaniti *et al.*'s case, it was found that the alignment of the peak may be influenced by the position of the accent relative to the word's right boundary and by the number of postaccidental syllables. However, the exact role of each of these two factors could not be discerned on the basis of their data. From their results it is concluded:

... that the canonical conditions which allow us to observe the full form of the Greek prenuclear accents involve at least two unaccented syllables following the accented one, preferably within the same word as the accent. If these conditions are met, then we observe a

very stable pattern of alignment of the H tone (Arvaniti *et al.*, 1998: p.22).

Similar prosodic effects on alignment were observed by Malliopoulos and Carayannis (1997). They carried out several experiments in order to determine characteristics of Greek intonation, necessary for the development of a phonetic implementation model for Greek intonation. The results of their experiments suggest that the onset of an accent-lending rise is invariably aligned with the start of the accented syllable, regardless of the number of preceding or following syllables and word boundary location. The only systematic exception to this alignment was observed in cases of "stress clash", when the immediately following syllable is also accented. In these cases, the peak (H\*) was shifted earlier, and the following peak was moved later. Furthermore, there also seemed to be an effect of "stress clash" on the sagging between successive high accents. However, Malliopoulos and Carayannis (1997) do not present results on this observation, and argue that more experiments are needed to justify this observation.

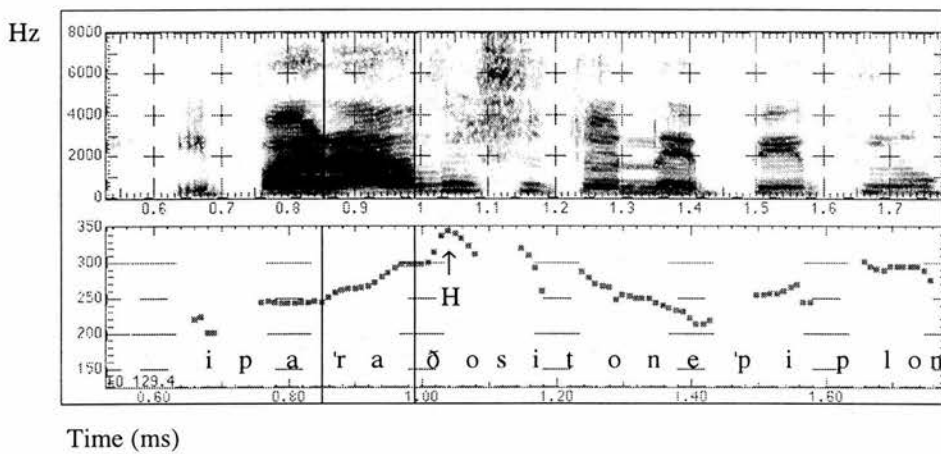


Figure 2.7. Spectrogram and F0 trace of the Greek pre-nuclear accent [i pa'radosi ton e'piplon] "The delivery of the furniture". The vertical lines delimit the stressed syllable of the word bearing the pre-nuclear accent.

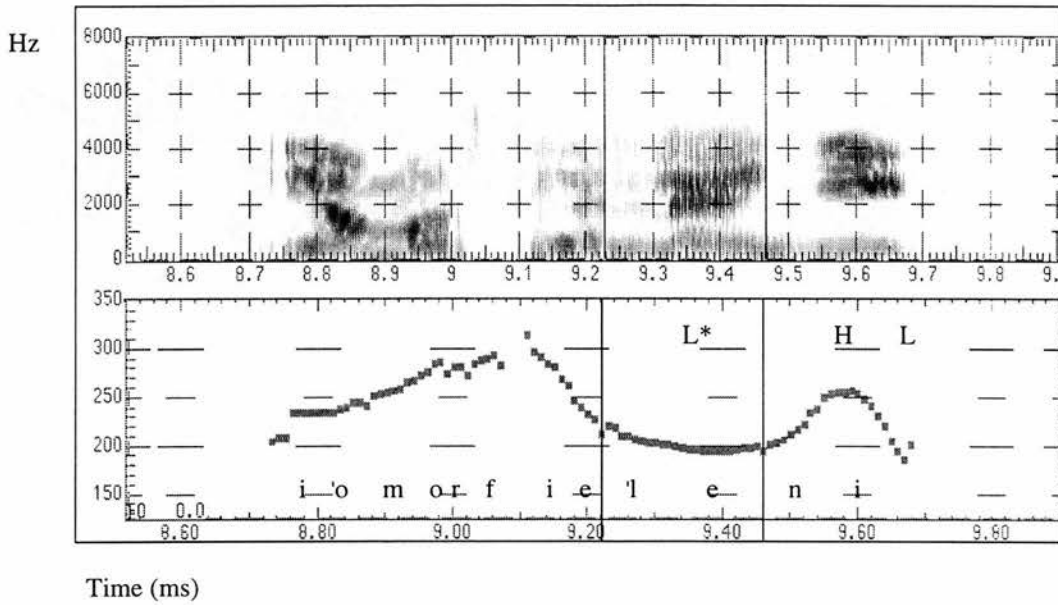
## 2.2. Greek yes/no questions

Greek yes/no questions are not marked by either syntactic or lexical means, and are distinguished only by intonation from otherwise identical statements. A characteristic feature of Greek yes/no questions is a rising-falling intonation pattern near the end of the utterance (Tsakirides, 1980; Waring, 1976, 1982; Arvaniti *et al.*, forthcoming). However, this is a rather sketchy description. In order to give a more exact description of Greek yes/no question intonation, it is necessary to take two factors into account: (i) the location of the most prominent word in the utterance, and (ii) the location of the lexical stress of the utterance-final word (Arvaniti, Ladd, and Mennen, forthcoming).

Consider the Dutch question *De mooie Helena?* "The beautiful Helen?". In this sentence, one could either put the main accent on *Helena*, and get a relatively neutral question, or one could emphasize the fact that she is beautiful by putting the main accent on *mooie* "beautiful". In the equivalent Greek question [i 'omorfi e'leni] "the beautiful Helen?" essentially the same thing can be done, where putting the main accent on [e'leni] will give a neutral question, putting it on ['omorfi] will emphasize her beauty. However, the actual phonetic realisation of this nucleus placement in Greek yes/no questions, is quite different than that in Dutch (or in English for that matter). In order to emphasize Helen's beauty, in Greek the F<sub>0</sub> has to remain low on ['omorfi] and the final rise-fall has to be on the stressed syllable of the final word [e'leni] "Helen". And, in order to emphasize [e'leni], the F<sub>0</sub> has to remain low into the vowel of ['le], and then rise and fall again at the end of the word. The above described contours can be seen in Figure 2.8, with the relatively neutral reading (i.e. the main accent on [e'leni]) in panel (a), and with emphasis on ['omorfi] in panel (b).

As can be seen in Figure 2.8, both contours show a rise-fall near the end of the utterance. However, according to Arvaniti *et al.* (forthcoming), Greek

(a)



(b)

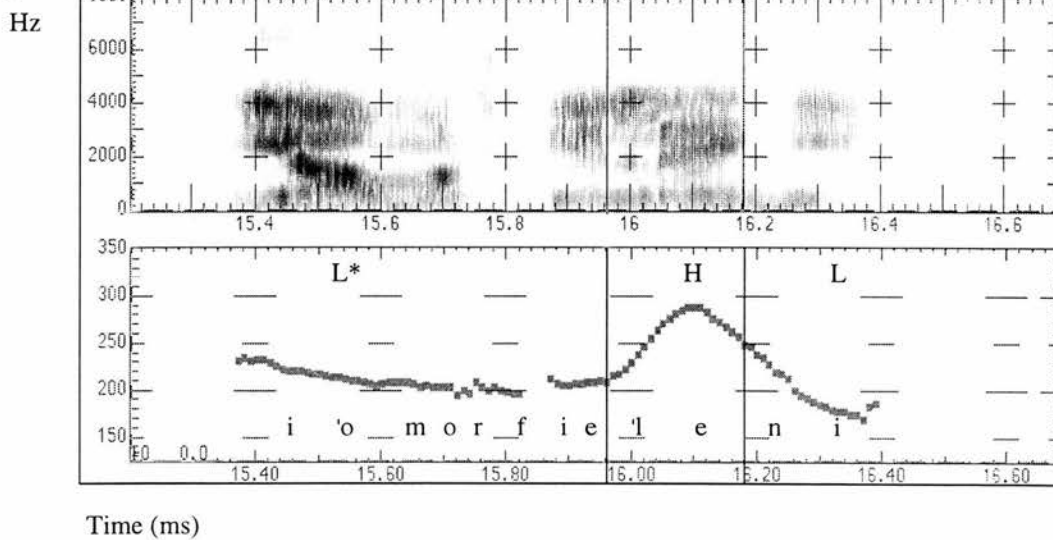


Figure 2.8. Spectrogram and F0 trace of the Greek yes/no question *[i'omorfieleni]* "The beautiful Helen?" In panel (a) the main accent is located on *[eleni]* "Helen", whereas in (b) it is located on *[omorfi]* "beautiful". The vertical lines delimit the beginning and end of the stressed syllable of the final content word.

listeners do not perceive this rise-fall as the nuclear accent. In yes/no questions like the one illustrated in panel (b) of Figure 2.8, it is the low-pitched word ['omorfi] which is perceived by Greek listeners as the nuclear accent, and not the final rise-fall on [e'leni]. Arvaniti *et al.* (forthcoming) analyse this contour in autosegmental-metrical terms as the sequence L\*HL. The L\* is said to associate with the nuclear syllable, whereas the HL sequence (the final rise-fall) is analysed as a boundary phenomenon. According to their analysis, the contours illustrated in Figure 2.8, are both variants of the same L\*HL sequence. Both contours consist of a L\* (which is associated with the main accent of the utterance), followed by a HL sequence (which is associated with the boundary). The difference is that in (b) the main accent is located early in the utterance and as a consequence nucleus and boundary phenomena are relatively far apart. In (a) we are dealing with the same sequence as in (b), but here the nucleus and the boundary are closer together, and all three tones of the L\*HL sequence are realised on just two syllables.

Inspection of Figure 2.8, however, also shows that there is another difference between the two contours. Although the H of the L\*HL sequence always occurs near the end of the sentence, it seems that the final H in contour (a) is later than in contour (b). Arvaniti *et al.* (forthcoming) show that this difference is due to the location of the main (nuclear) accent. If the nuclear accent is on the last word (as in panel b), the H occurs on the last syllable of the sentence. If, on the other hand, the nuclear accent is on an earlier word (as in panel a), then the H occurs on the stressed syllable of the last word (Arvaniti *et al.*, forthcoming). From their results it can be concluded that for an accurate description of Greek yes/no question intonation, both the location of the nuclear accent as well as the location of the lexical stress of the utterance-final word, need to be taken into account.

### 3. Comparison of Dutch and Greek intonation

#### 3.1. An autosegmental-metrical analysis of Dutch and Greek intonation

In the previous sections intonational analyses of prenuclear accents and yes/no questions in Dutch and Greek were presented, in order to make a comparison of the intonation patterns across these two languages. However, on the basis of the analyses provided in the previous sections such a comparison is difficult, as the analyses are (to an extent) based on different frameworks. In order to come to any reliable conclusions about the significance of differences and/or similarities between the Dutch and Greek intonation patterns it is necessary to agree on a common framework. Therefore, in this section, the analyses presented in the previous sections will be reanalysed in terms of the standard autosegmental-metrical framework of intonational analysis, as exemplified by Pierrehumbert (1980) and Pierrehumbert and Beckman (1988).

As some readers may not be familiar with this framework, the notational conventions employed in this framework will be presented first. In the standard autosegmental-metrical approach to the study of intonation three types of tonal categories are recognised: *pitch accents*, *phrase accents*, and *boundary tones* (as described in Chapter 1, section 5.2). Pitch accents can consist of a single high or low tone (marked H\* or L\*) or a combination of two tones. An asterisk is used to indicate the central tone of a pitch accent, and this starred tone can be preceded or followed by another tone, called respectively a 'trailing' or 'leading' tone. The trailing or leading tone is marked with a raised hyphen (H<sup>-</sup> or L<sup>-</sup>). The bitonal accents are joined with a plus sign (e.g. L+H\*, or L\*+H). It is assumed that the central or starred tone of bitonal accents is associated with the main

stressed syllable, while the unstarred tone remains unassociated (Beckman and Pierrehumbert, 1986).<sup>8</sup>

Dutch prenuclear accents are characterised by a fairly steep rise followed by a more gradual fall. In terms of the standard autosegmental-metrical framework it would have to be analysed as a H\* accent. In the same framework Greek prenuclear accents would also be analysed as instances of a H\* accent.<sup>9</sup>

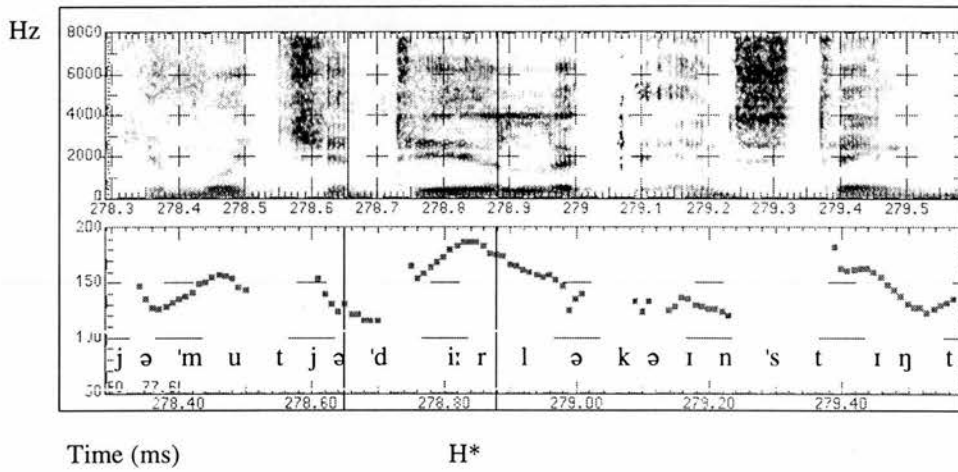
In other words, Greek and Dutch prenuclear accents are similar in the sense that they both have a rise that begins near the onset of the accented syllable, followed by a gradual fall that usually extends over several unstressed syllables. However, in Dutch prenuclear accents the rise reaches its peak somewhere *within* the accented syllable, whereas in Greek it peaks in the *following vowel*. Thus, although Greek and Dutch prenuclear accents arguably share the same phonological analysis, they differ in phonetic detail. That is, the alignment of the peak with respect to the segmental string is different for the two languages, with the H\* earlier in Dutch than in Greek prenuclear accents. In Figure 2.9, modified from Figure 2.4, the intonation patterns of Dutch and Greek prenuclear accents are shown, together with their analysis in autosegmental-metrical terms.

---

<sup>8</sup> Recently, it has been argued that the notion of starredness is not without problems and that a clear(er) definition is needed (Arvaniti, Ladd, and Mennen, in press; Ladd, 1998). However, this will not be discussed in detail here as this is largely beyond the scope of the present thesis.

<sup>9</sup> Arvaniti, et al. (in press) discuss the difficulty of deciding the correct autosegmental analysis of Greek prenuclear accents. This difficulty arises from the fact that Greek prenuclear accents consist of a L and H tone, of which the L is aligned just before the accented syllable and the H is aligned just after the accented syllable. That is, none of the tones is aligned in time with the accented syllable, although both tones seem to be associated with the accent. Although the analysis of Greek prenuclear accents is inconclusive, it seems clear that the H tone is the perceptually more salient tone. They argue that "if starredness is at all related to strength, intuitively one would expect the perceptually more salient tone to be the starred one."

(a)



(b)

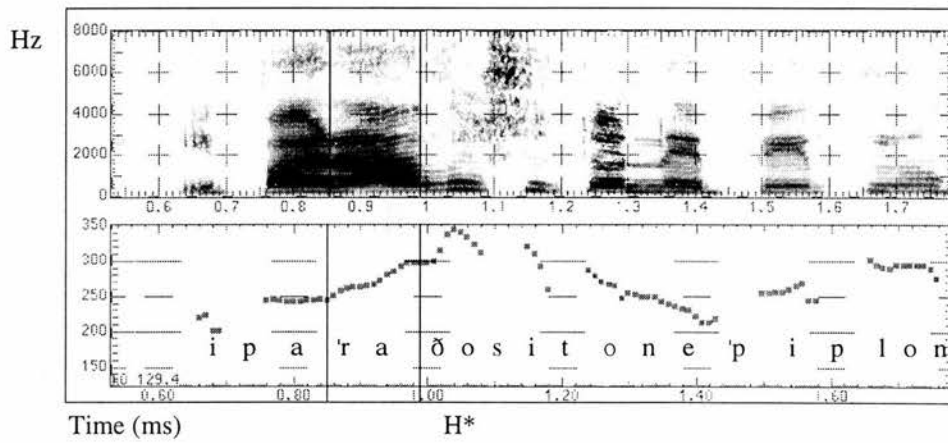
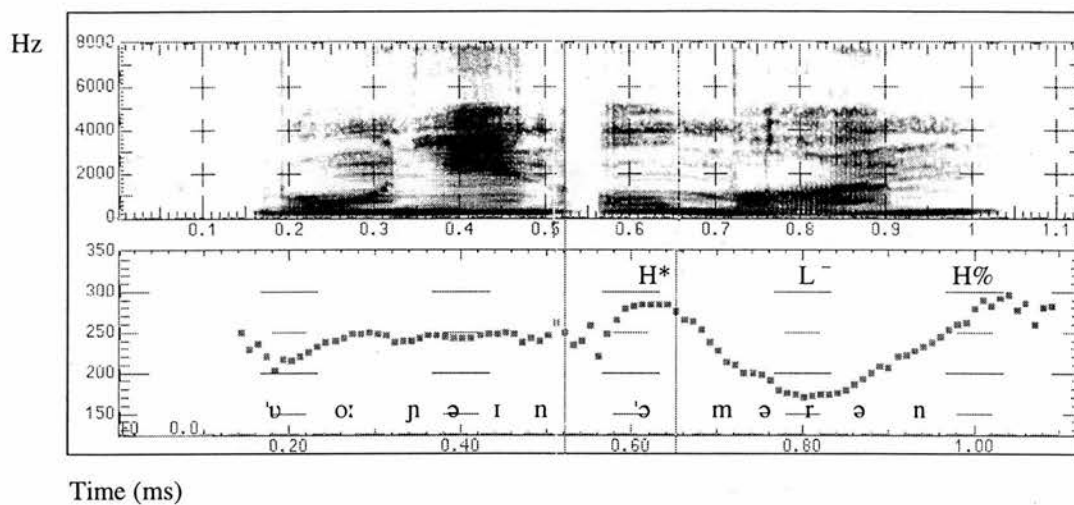


Figure 2.9. Spectrogram and F0 trace of a Dutch (a) and Greek (b) pre-nuclear accent. The vertical lines delimit the stressed syllable of the words bearing the pre-nuclear accent. Both accents are analysed as H\* accents, but it is clear that the H is earlier in Dutch than in Greek.

Yes/no question intonation, on the other hand, exhibits more differences between Dutch and Greek. Most importantly, Greek only has *one* yes/no question pattern, whereas Dutch has *several* patterns. In autosegmental-metrical terms Greek yes/no questions are analysed as L\*H-L%, that is a L\* nuclear accent followed by a H phrase accent and a L% boundary tone (Arvaniti *et al.*, forthcoming). Dutch yes/no questions, on the other hand, are of several types, with either a H\* or a L\* nuclear accent. In autosegmental-metrical terms they are analysed as (i) H\* L- H% (i.e. a H\* nuclear accent followed by a low phrase tone and a H% boundary tone), or (ii) L\* H- H% (a L\* nuclear accent followed by a H phrase tone and a H% boundary tone). Recently, Gussenhoven and Rietveld (1997; forthcoming) proposed that a Dutch question intonation contour can be realised as a low or a high rise. They argue that “these two contours can either be seen as the extremes on a continuum of phonetically different realisations of the same phonological contour L\* H- H%, or as phonetic realisations of two phonologically different contours, L\* H- H% [‘low rise’] and H\* H% [‘high rise’]” (Gussenhoven and Rietveld, 1997: p.169). The results of their experiments suggest that the low rise and the high rise may be categorically distinct contours of Dutch. Haan (personal communication) recognises two main yes/no question intonation contours, H\* L- H%, and L\* H- H%, both of which can have ‘boundary linking’. In contours where there is boundary linking, the L or H phrase tone gets deleted and the boundary tone immediately follows the H\* or L\*. Thus H\* L- H% becomes H\* H% (i.e. Gussenhoven and Rietveld’s ‘high rise’), and L\* H- H% becomes L\* H%. In Haan’s corpus of 400 yes/no and declarative questions, the H\* L- H% contour was the most frequent and seems to be unmarked. The H\*H% contour was less frequent and possibly expresses more surprise. When the nuclear accent was low, the preferred contour seemed to be L\*H% rather than L\* H- H%, which occurred less frequently.

(a)



(b)

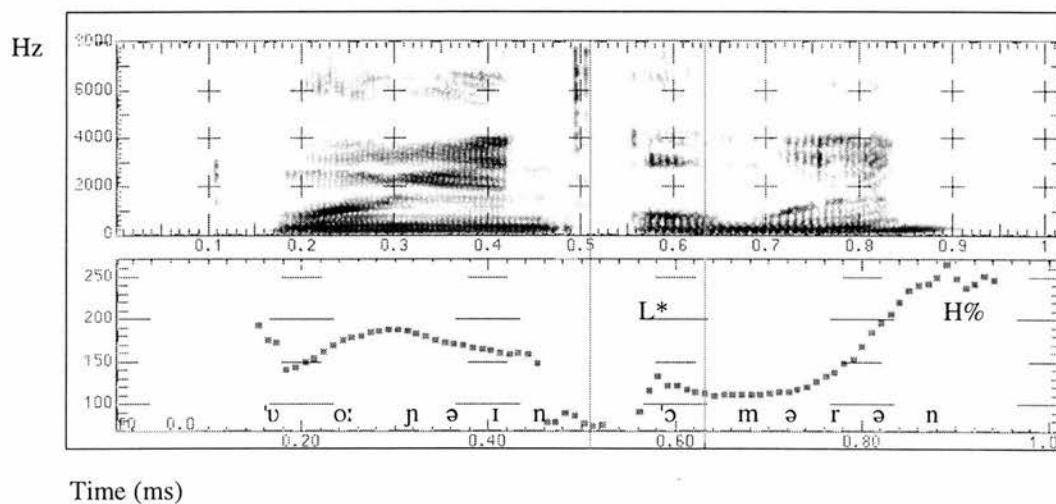
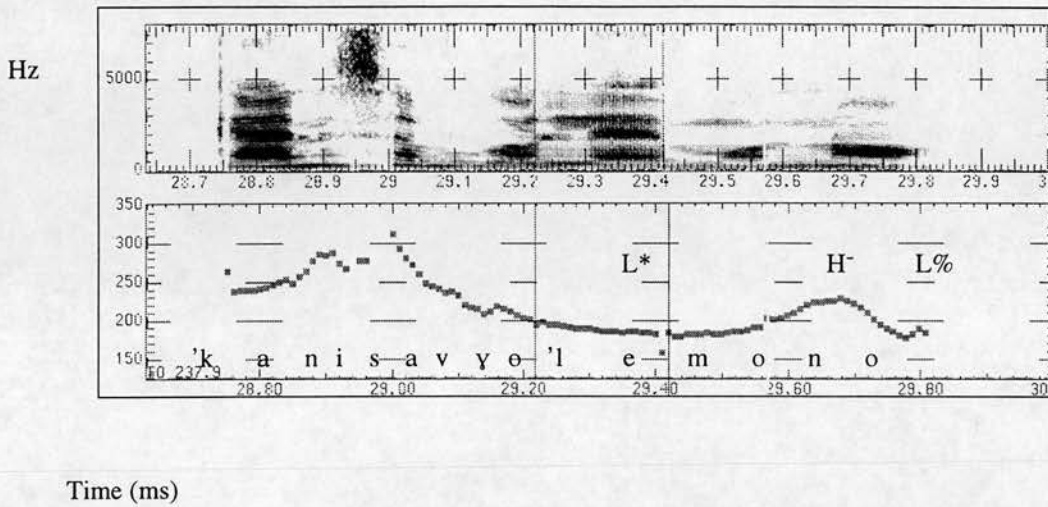


Figure 2.10. Spectrogram and F0 trace of the Dutch yes/no question "Woon je in Ommeren?" "Do you live in Ommeren?", together with their autosegmental analyses. Panel (a) shows an example of a H\*L-H% contour. Panel (b) shows an example of a L\*H% contour. The vertical lines delimit the stressed syllable of the nuclear word.

(a)



(b)

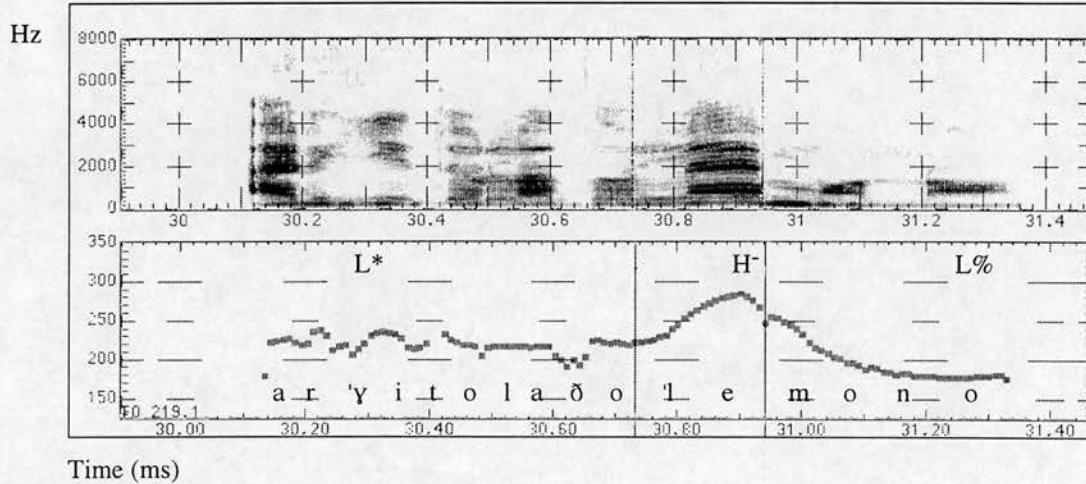


Figure 2.11. Spectrogram and F0 trace of the Greek yes/no questions: (a) [kanis avyo'lemono] "Are you making egg-and-lemon sauce?", and (b) [ar'yi to laðo'lemono] "Does the oil-and-lemon sauce take long?". In panel (a) the main accent is located on [avyo'lemono] "egg-and-lemon sauce", whereas in (b) it is located on [ar'yi] "does take long". The autosegmental analysis is given above the contours. The vertical lines delimit the beginning and end of the stressed syllable of the final content word.

Thus, Dutch seems to have at least two yes/no question contours, of which the  $H^* L^- H\%$  and the  $L^* H\%$  contours seem to be the most common. In any case, there are more question contours in Dutch than in Greek, and all Dutch yes/no questions have different tunes than the Greek yes/no question contour. In Figure 2.10 (adapted from Figure 2.5) the most common Dutch yes/no question contours are shown, together with their autosegmental analyses. Examples of Greek yes/no questions are shown in Figure 2.11.

Besides the fact that Greek only has one yes/no question pattern and Dutch has at least two patterns, it is clear that the Greek and Dutch questions have different tunes. One obvious example of this difference in tunes is that the Dutch questions end in a rise, whereas the Greek questions end in a fall, as indicated by the  $H\%$  and  $L\%$  boundary tones.

There are, however, some differences which are not immediately clear from the autosegmental-metrical analyses. First of all, as mentioned in the previous subsection, in Greek yes/no questions the position of the peak is affected by the location of the main or nuclear accent, something which does not happen in Dutch. Specifically, it seems to be the case that the  $H^-$  occurs on the utterance-final syllable if the nuclear accent is on the last word. If, on the other hand, the nuclear syllable is on an earlier word, the  $H^-$  occurs on the lexically stressed syllable of the last word (Arvaniti et al., forthcoming). There is no obvious counterpart of this H phrase accent association in Dutch.

A second point of difference between Greek and Dutch yes/no questions, which is not immediately obvious from the analyses provided (but may nevertheless be important), is their prominence pattern. In Greek yes/no questions the default location for the nuclear accent is the verb (just as in Hungarian and Romanian, cf. Ladd, 1996). That is, when the greatest prominence is on the verb, the question is relatively neutral, as in the following example:

θes tsi'ɣaro? Would you like a cigarette (!)

In Dutch, on the other hand, having the main accent on the verb would make the question non-neutral.

ɔɪ jə ən siɣa'ret? Would you like a cigarette?

and the interpretation would shift from a neutral question to something like 'Come on, make up your mind. Would you like a cigaret or not?

A further difference between Greek and Dutch is concerned with the association of rise-falls. Although rise-falls may occur in both languages, in Dutch they usually occur in statements rather than questions, and always associate with a lexically stressed syllable. That is, a rise-fall in Dutch only occurs *on* or *before* the nuclear accent. In Greek, however, the peak of the final rise-fall occurs *after* the nuclear accent, whereas in Dutch it can only be nuclear or prenuclear. It thus becomes clear that Greek and Dutch yes/no questions differ in their phonological structure.

Summarised, the following similarities and differences between Dutch and Greek intonation can be found:

- Dutch and Greek prenuclear accents share the same phonological analysis, i.e. H\*
- Dutch and Greek prenuclear accents differ in phonetic detail: the H\* is earlier in Dutch than in Greek prenuclear accents
- Greek has only one yes/no question pattern (L\*H-L%), whereas Dutch has at least two patterns, i.e. H\*(L-)H%, or L\* (H-)H%.
- the nuclear accent in Greek yes/no question intonation is L\*, whereas in Dutch it can be either L\* or H\*

- the peak of the rise-fall in Greek yes/no questions occurs *after* the nuclear accent, whereas in Dutch a rise-fall can only be nuclear or prenuclear.
- in Greek yes/no question intonation the H<sup>-</sup> phrase accent is affected by the location of the nucleus: if the nuclear accent is on the last word, the H<sup>-</sup> occurs on the sentence's last syllable; if it is on an earlier word, the H<sup>-</sup> occurs on the stressed syllable of the last word. There is no obvious counterpart of the H phrase accent alignment in Dutch
- in all Greek yes/no questions the H<sup>-</sup> is followed by a L% boundary tone. In Dutch yes/no questions the boundary tone is usually H%
- the neutral location for the nuclear accent in Greek yes/no questions is the verb, whereas in Dutch having the main accent on the verb would make the question non-neutral

### 3.2. Flege's notions of 'new' and 'similar'

In Chapter 1 a current L2 speech model was presented, the Speech Learning Model (SLM), which was developed to account for segmental aspects of second language learning. Although it is accepted that prosodic factors are an important source of foreign accent, no attempt has been made to expand this model to account for nonsegmental aspects of L2 learning (Flege, 1995; Flege *et al.*, 1995). In this section an attempt will be made to apply the assumptions this model makes to the intonation of native (Dutch) speakers of L2 Greek.

The SLM is based on the 'phonetic similarity' between L1 and L2 sounds (Flege, 1995; Flege *et al.*, 1995). It is argued that when sounds are highly dissimilar (called 'new' sounds), listeners will eventually (after some initial difficulty) form distinct categories for these 'new' L2 sounds. As a result, perception and production of these sounds will usually be

good. Sounds which are similar to but not identical with L1 sounds (called 'similar' sounds) are the most problematic for L2 learners. These 'similar' sounds will be classified according to L1 categories even after considerable experience with the L2. This will result in phonetically inaccurate production of these sounds and difficulty in perceiving that these sounds differ from the corresponding L1 sound.

According to the SLM, learners relate L1 and L2 sounds perceptually at a 'position sensitive allophonic level' (Flege, 1995:p.239). In other words, the SLM takes context-dependent phonetic segments as the appropriate level of analysis, so that the syllabic position, the phonological context, and the allophonic variations of the segments are taken into account when determining the similarity between L1 and L2 sounds.

In the case of prenuclear accents, it is quite straightforward to determine the degree of similarity between Greek and Dutch intonation patterns. Both Dutch and Greek prenuclear accents can be analysed as instances of a H\* accent. That is, in Flege's terminology the prenuclear Greek accent is seen as 'similar' to the Dutch prenuclear accent. However, they are not the same, as they differ in phonetic detail (as can be seen in Figure 2.9). The SLM would consequently predict that non-native (Dutch) speakers of Greek would initially fail to recognise that the differences between the L1 and L2 H\* patterns are phonetically relevant. As a result, category formation would be blocked, and the Dutch learners of Greek would classify the similar L2 pattern according to their L1 category. This would lead to inaccurate production of the L2 pattern. The SLM holds that inaccurate production would persist even after long exposure to the L2. It is further predicted that, over time, the L2 learners will merge the properties of the L1 and L2 patterns. That is, Dutch learners of Greek would develop a 'merged' system, intermediate between the L1 and the L2 norm. As in this case the difference between Dutch and Greek prenuclear H\* accents lies in the alignment of the peak, the SLM would predict that Dutch learners of Greek would align the peak in Greek H\* accents somewhere between the Dutch and the Greek norm.

Applying the notions of 'new' and 'similar' to differences in Greek and Dutch yes/no question intonation is less straightforward. First of all, Greek and Dutch have different yes/no question tunes: Greek has a L\*H<sup>-</sup>L% contour, whereas Dutch yes/no questions are either H\*(L<sup>-</sup>)H%, or L\*(H<sup>-</sup>)H%. In other words, the Greek yes/no question tune is highly dissimilar from the Dutch yes/no question tunes. It is not clear how this difference should be interpreted in terms of the SLM, as this difference has to do with the meaning and function of intonation, rather than with an acoustic-phonetic dissimilarity. That is, there are rises and falls in both Dutch and Greek. It is just that the specific tunes used in Greek yes/no questions are not used to express interrogativity in Dutch. The SLM is not concerned with this kind of dissimilarities between the L1 and L2, and therefore on the basis of this difference in question tunes it cannot be determined whether the L2 question pattern is 'new' or 'similar'.

Another difference is the fact that the neutral nucleus location in Greek yes/no questions is on the verb, whereas in Dutch having the main accent on the verb would make the question non-neutral. Again, it is not clear how this should be seen in terms of the SLM, but it is nevertheless a difference which may have to be considered in making predictions about the difficulty Dutch learners may experience with production and perception of Greek yes/no questions.

On the other hand, the phonetic shape of Greek yes/no questions is not that dissimilar from Dutch intonation patterns, in the sense that a rise-fall is very common in Dutch. However, in Dutch a rise-fall would usually occur in statements rather than yes/no questions. Furthermore, in Dutch a rise-fall always associates with a lexically stressed syllable. In Greek, on the other hand, the location of the rise-fall is affected by the location of the nucleus: if the nuclear accent is on the last word, the rise-fall occurs on the sentence's last syllable (as shown in the top panel of Figure 2.11); if it is on an earlier word, the rise-fall occurs on the stressed syllable of the last word (as shown in the bottom panel of Figure 2.11).

That is, there is a difference in the location of the peak in nucleus-final (NF) and nucleus-non-final (NNF) yes/no questions. The latter contour is thus more similar to Dutch, in the sense that the rise-fall occurs on a lexically stressed syllable. A rise-fall on an unstressed syllable, as can occur in Greek NF yes/no questions, is not possible in Dutch. Therefore, on the basis of the phonetic shape, the SLM would consider the NNF yes/no question as 'similar', and the NF yes/no question as 'new'.

A related difference is the fact that a rise-fall in Dutch is always prominence-lending, whereas in Greek yes/no questions it is not. In Greek yes/no questions, the nuclear accent is L\*, and the rise-fall occurs *after* the nuclear accent. Since a rise-fall in Dutch is prominence-lending (and thus is either prenuclear or nuclear), it is likely that Dutch listeners will perceive the syllable bearing the final rise-fall as the most prominent syllable in Greek yes/no questions. So, regardless of the location of the nucleus (NF or NNF) it is likely that it is the utterance-final word which will be perceived as the word bearing the nuclear accent. For Greek speakers, however, it is the low-pitched lexically stressed syllable which is perceived as the most prominent in the utterance (Arvaniti *et al.*, forthcoming).

From the above discussion it can be inferred that the (dis)similarity between Dutch and Greek yes/no questions cannot readily be captured in terms of Flege's SLM. It seems that, besides phonetic similarity, there may be additional factors which determine the degree of difficulty Dutch learners of Greek experience when producing Greek yes/no questions. However, since one of the aims of this thesis is to evaluate the efficacy of the SLM in explaining L2 prosodic data, predictions will only be made on the basis of phonetic similarity, and any other factors will be left aside for later discussion. On the basis of phonetic similarity, one has to conclude that the perceptual distance of Dutch and Greek yes/no questions may not be the same for the two different focus readings of Greek yes/no questions, i.e. nucleus-final (NF) or nucleus-non-final (NNF). It is likely that the perceptual distance is greater for NF than for NNF yes/no

questions, as in the former the final rise-fall occurs on an unstressed syllable, something which is not possible in Dutch. In Dutch a rise-fall is always associated with a lexically stressed syllable. Therefore, the Greek NNF pattern, with its rise-fall on the lexically stressed syllable of the utterance-final word is more similar to Dutch than the NF pattern (where the rise-fall can occur on an unstressed syllable). Therefore, the NF pattern should be seen as 'new', whereas the NNF should be seen as 'similar'. According to the SLM, perception and production of a new sound should usually be unproblematic, since learners will be able to establish new categories for new L2 sounds. On the other hand, sounds which are similar should be problematic for L2 learners, and will be classified according to L1 categories. If applied to Greek yes/no question intonation, the SLM predicts that Dutch learners of Greek should not experience any problem producing the 'new' NF contour, but production of the 'similar' NNF contour should be inaccurate (even after considerable experience with the L2).

## Chapter 3

### Prenuclear accents

#### 1. Introduction

In this chapter the production of Greek prenuclear accents by non-native (Dutch) speakers of Greek is examined. In Chapter 2 it was seen that Greek and Dutch prenuclear accents are both analysed as instances of a H\* accent, but that they differ in phonetic detail. More specifically, it was suggested that they differ in the alignment of the peak (H), which seems to be earlier in Dutch than in Greek. However, so far there are no data available on the precise alignment of the peak (H) or the preceding valley (L) in Dutch prenuclear accents. Therefore, the first experiment (experiment 1) reported in this chapter is a pre-test to establish the position of the L and H in Dutch prenuclear accents. The next two experiments (experiments 2 and 3) were run to test the core issues of this chapter, namely (i) whether non-native (Dutch) speakers of Greek will acquire native-like alignment of prenuclear Greek accents, (ii) whether the acquisition of an L2 (Greek) has any effect on the alignment of these accents in the L1 (Dutch), and (iii) whether the Speech Learning Model (SLM) can explain the alignment pattern (in the L2 *and* L1) of the non-native (Dutch) speakers of Greek.

In first instance, it seemed that a comparison of the alignment in Dutch and Greek prenuclear accents, was to be quite straightforward. However, visual inspection of some prenuclear Dutch intonation patterns, suggested that Dutch alignment patterns are influenced by the phonological length of the lexically stressed syllable: the peak (H) seems to occur on the lexically stressed vowel when that vowel is phonologically long, but it occurs in the following consonant when the stressed vowel is

phonologically short.<sup>1</sup> This observation made the intended comparisons much more complicated, as such alignment differences are not found in Greek prenuclear accents.

Table 3.1. *Experimental design for each of the experiments described in this chapter..*

Experiments	Group	Materials	Comparisons
Exp. 1	Group D	Set D-long Set D-short	Group D (D-long) vs. Group D (D-short)
Exp. 2	Group D	Set D-long Set D-short	a) Group D (D-long) vs. Group DG (Greek), and Group DG (Greek) vs. Group G (Greek)
	Group DG	Set Greek	b) Group D (D-short) vs. Group DG (Greek), and Group DG (Greek) vs. Group G (Greek)
	Group G	Set Greek	
Exp. 3	Group D	Set D-long Set D-short	a) D (D-long) vs. DG (D-long) b) D (D-short) vs. DG (D-short)
	Group DG	idem	
Extra analysis	Group DG	Set D-long Set D-short Set Greek	DG (D-long) vs. DG (D-short) vs. DG (Greek)

In order to investigate the aforementioned issues, three sets of sentences were used with three groups of subjects. The sets were: a Dutch set with phonologically long vowels in the stressed syllable of the test word ("Set D-long"), a Dutch set with phonologically short vowels in the stressed syllable of the test word ("Set D-short"), and a set with Greek prenuclear accents ("Set Greek"). The three groups of speakers were: a

<sup>1</sup> This pattern was observed in a pilot experiment designed to replicate an experiment on Greek prenuclear accents (the Greek experiment is described in Arvaniti, Ladd, and Mennen, 1998). The pilot experiment consisted of 32 sentences with different vowels in the

group of native Dutch speakers ("Group D"), a group of native Greek speakers ("Group G"), and a group of non-native (Dutch) speakers of Greek ("Group DG"). Group D read the Dutch sets (Set D-short and D-long) only, Group G read the Greek set (Set Greek) only, and Group DG read both the Dutch sets (D-short and D-long) as well as the Greek set (Set Greek). A more detailed description of the sets of materials, groups of speakers, and comparisons will be given in the method section of each experiment. Table 3.1. summarises the groups and sets of materials used, and the comparisons which were made.

## 2. Experiment 1

### 2.1. Introduction

In Chapter 2 it was suggested that Dutch and Greek native patterns of alignment of the peak (H) in prenuclear accents are different. In Greek the peak is very stable (at least in words with lexical stress on the antepenultimate syllable) and is consistently aligned around 10 to 20 ms after the onset of the first postaccentual vowel. Figure 3.1. shows an example of a Greek prenuclear accent (repeated from Figure 2.7). In Dutch, on the other hand, it was suggested that the peak is aligned earlier, somewhere *within* the accented syllable. As was previously described, results of a pilot experiment suggested that the alignment of the peak in Dutch prenuclear accents is affected by the phonological length of the lexically stressed vowel. Experiment 1 was designed to verify whether this impressionistic observation that Dutch prenuclear accents exhibit two modes of alignment of the peak is indeed true, and consequently to establish whether there are cross-linguistic differences in alignment between Greek and Dutch.

---

accented syllable of the test word, and was recorded with four native speakers of Dutch. The pattern was observed in the data of all four speakers.

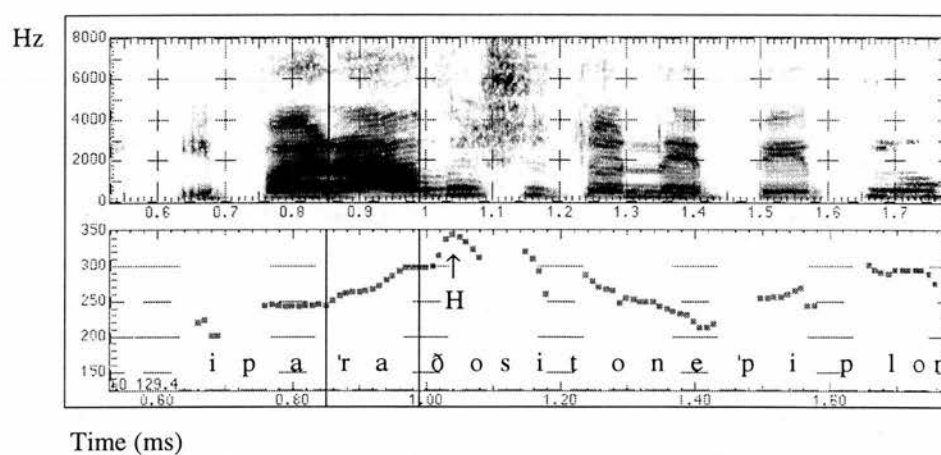


Figure 3.1. Spectrogram and F0 trace of the Greek prenuclear accent [i pa'radosi ton e'piplon] "The delivery of the furniture". The vertical lines delimit the stressed syllable of the word bearing the prenuclear accent.

## 2.2. Method

### Materials

The items were 40 Dutch test sentences, which were of two types:

A. with phonologically long vowels in the accented syllable of the test word (i.e. the word bearing the prenuclear accent). This set will be referred to as “set D-long”.

B. with phonologically short vowels in the accented syllable of the test word. This set will be referred to as “set D-short”.

Dutch has a 16 vowel inventory with both phonologically long and short vowels, as well as diphthongs. In Table 3.2 a classification of the sixteen Dutch vowels is given (from Booij, 1995).

Table 3.2. *The Dutch vowels (from Booij, 1995)*

Short vowels	ɪ, ɛ, ɔ, ʏ, ɑ
Long vowels	i, e, o, y, a, ø, u
Schwa	ə
Diphthongs	ɛɪ, œy, ɔu

Condition A was designed with 5 phonologically long vowels (/i, e, a, o, and y/) in the accented syllable of the test words. Condition B was constructed with the short counterparts of the vowels which appeared in the test words of condition A (i.e. /ɪ, ɛ, ɑ, ɔ, and ʏ/). In other words, the vowels of Set D-long and Set D-short were paired, with the phonologically long vowels in set D-long, and the short vowels in set D-short. The length distinction here used is a phonological distinction (i.e. long vowels behave as two units, whereas short vowels behave as one unit), and is not necessarily reflected in a phonetic distinction. For example, the

phonologically high long vowels /i, y, u/ have the same average duration as the short vowels /ɪ, ɛ, ə, ʏ, ʌ/ (cf. Booij, 1995: p. 4-5).

In each condition the test words were exclusively words with lexical stress on the antepenultimate syllable and were followed by two to five unaccented syllables. This was done in order to avoid effects of prosodic context on the alignment of the peak (as described in Chapter 1, section 5.3.1, and Chapter 2, section 1.3), such as those found by for instance Prieto (1995), Silverman and Pierrehumbert (1990), and Arvaniti *et al.* (1998). In order to avoid difficulties with the peak alignment measurements, test words were chosen with only sonorants in the relevant syllables (i.e. the accented syllable and the following consonant(s)), so that the F0 contour would be uninterrupted. The materials used in this experiment were also intended to be used in further experiments (experiment 2 and 3), and one of the aims was to test whether there are cross-linguistic differences between Dutch and Greek prenuclear accents. Therefore, the Dutch materials were designed to match a Greek set (which will be described below in experiment 2), as closely as possible, allowing future comparisons between Dutch and Greek. As far as possible, the materials were constructed with simple consonant-vowel (CV) sequences in the accented syllables of the test words. However, a small number of accented syllables with a more complex structure (either with a syllable-initial consonant cluster, or with heterosyllabic consonants following the accented vowel) were included. This was done in order to match the design of the aforementioned Greek set, which included a small number of accented syllables with a more complex structure. On the basis of previous research (Arvaniti *et al.*, 1998), it was not expected that initial consonant clusters (such as [ˈdralɔndə] “lingering”) nor heterosyllabic consonants (such as [ˈmɔndəxɛ] “emancipated”) would have an effect on the alignment of the peak. In any case, the number of test words with heterosyllabic consonants or syllable-initial consonants was small (i.e. 15% of all the test words for the former, and 10% for the latter). Examples

of the items included in this experiment are given in Table 3.3. A full list of test items can be found in Appendix A.

Table 3.3. *Experiment 1: sample test items for each of the two sets (D-long, and D-short); the test words are underlined.*

<u>D-long:</u>	[hɛi kɔn də 'ma:lɔndə xə'dɔxtə 'ni:t œytət 'hɔ:ft setə]
	(He could the persistent thoughts not out the head put)
	"He could not get the <u>persistent</u> thoughts out of his mind."
	[zə kɔndə də 'li:ri:sə 'steil fən zɛin rɔ:'mɛns 'ni:t vɑ:rdɛ:rə]
	(They could the lyrical style of his novels not appreciate)
	"They could not appreciate the <u>lyrical</u> style of his novels."
	[jə mu:t in 'rɔ:məxə 'sɔuzə ən sxø:tjə kə'ŋɑk du:n dət is 'lɛkər]
	(You should in creamy sauces a dash cognac put, that's nice)
	"You should put a dash of cognac in <u>creamy</u> sauces, that's nice."
<u>D-short:</u>	[ɪk vɔs də 'drɛmɛndə tɛ:lə'fɔ:ntjəs fən dət fər'fe:lɔndə mɛns 'spɪ:xst]
	(I was the pestering phone-calls of that annoying woman sick-of)
	"I was sick of the pestering phone-calls of that annoying woman."
	[mɛt ha:r bə'mɪnələk xə'drɑx kən zə 'i:dərə:n ɔm ha:r 'fɪŋər vɪndə]
	(With her charming behaviour can she anyone round her finger twist)
	"With her <u>charming</u> behaviour she could twist anyone round her finger."
	[ɪkɔn mən 'mɔrɛndə kɔ'le:ɪxɑ: ni:t mɛr 'lɪxtə of 'sɪn]
	(I could my murmuring colleague not anymore stand)
	"I could not stand my <u>murmuring</u> colleague anymore."

## Subjects

Three female (D1, D4 and D5) and two male subjects (D2 and D3) participated in this experiment in individual sessions. All speakers were native speakers of Dutch, and were in their thirties and forties. The speakers were educated, and had no self reported hearing or reading disabilities. In the remainder of this thesis the group of native Dutch speakers will be referred to as Group D.

It was originally intended to select only monolingual speakers for this experiment, as proficiency in another language may have an effect on the L1 (as previously described in Chapter 1, section 3.4). However, this proved to be an unrealistic criterion for speaker selection, especially among educated speakers. Therefore, all speakers were also reasonably competent in English. It was hypothesised that the effect on the L1 would be more obvious in more advanced speakers of a second language. Therefore, care was taken not to include very advanced or near-native speakers of a second language.<sup>2</sup>

### Procedure

Prior to the recording the speakers were instructed to read the sentences as naturally as possible, and were asked to repeat any misread sentence. The experimenter monitored whether subjects read the sentence as it was printed, and were not omitting, adding or altering any words. If subjects misread a sentence, they were asked to repeat it.

The speakers read each test sentence twice from a randomized set of cards, each card containing one typed sentence. In order to avoid listing effects, the cards were interspersed with cards which contained material for other (pilot) experiments, which will not be reported here. The entire recording session took approximately 15 minutes.

### Apparatus and measurements

All materials were recorded on professional equipment at different locations: in the studio of the Department of Phonetics, University of Amsterdam, in the studio of the Department of Linguistics, University of

---

<sup>2</sup> In fact, it is not easy to define what it means to be near-native. The term itself, “near-native”, implies that it is almost the same as native, but somehow falls short (Sorace, 1993: p. 23). Since this thesis is concerned with L2 acquisition of *prosody*, other aspects of L2 acquisition (i.e. syntax, semantics) are not of immediate interest here. Therefore, the

Groningen, or at a quiet room in the speakers' home. The speakers were not paid for their participation.

The recordings were made on digital audio tape (DAT). The test materials were digitised at a 16kHz sampling rate with appropriate low-pass prefiltering. The present author selected the first acceptable repetition for further measurement using the following criteria for selection: A repetition was considered unacceptable when it contained a disfluency or when the recording was noisy, when the test word was immediately followed by a phrase boundary, or when an otherwise different intonation contour from the one anticipated was produced (e.g. when the sentence was read as a question instead of a statement). All other sentences were considered acceptable.

The selected sentences (i.e. 40 sentences for each speaker) were analysed using a Sun SPARC workstation with ESPS Waves<sup>+</sup> speech analysis facilities. Pitch tracks were obtained using the pitch-tracking facility of Waves<sup>+</sup>, with a 49 ms cos.<sup>4</sup> window moving in 10 ms steps. Durational measurements were made by marking several points from waveforms in combination with wide-band spectrograms and F0-tracks. The following points were marked: the consonant onset of the accented syllable of the test word (C0), the end of the accented vowel (C1)<sup>3</sup>, the peak (H) and the valley preceding the peak (L).

In general, the peak (H) was easy to find, and was defined as the highest F0 point around the end of the accented syllable of the test word. No attempt was made to compensate for possible microprosodic effects on F0, i.e. the absolute F0 maximum was consistently chosen. The alignment of H was expressed relative to the end of the accented vowel (C1). The L was defined as the absolute F0 minimum in the location of the accented syllable of the test word. Just as for the location of the H, no attempt was

---

term "near-native" is used to refer to speakers who could be mistaken for a native speaker of the L2, when judged on pronunciation only.

<sup>3</sup> The end of the accented vowel of the test word was in most cases also the end of the accented syllable. However, this was not always the case as in a small number of cases the accented vowel was followed by heterosyllabic or ambisyllabic consonants.

made to compensate for any segmental effects on F0. The alignment of L was expressed relative to the onset of the accented syllable (C0). The relevant segments (C0, and C1) were marked using visual information from waveforms in combination with wide-band spectrograms and F0-tracks, following the criteria for segmentation described in (Van Zanten, Damen, and Van Houten, 1991). Changes in amplitude and period structure were used as indicators of segment boundaries. The boundaries between vowels and nasals or laterals were marked at the point where sudden changes in both amplitude and period structure occurred. In case there were no sudden changes (which were sometimes not observed, most often at the boundaries between vowels and laterals) the boundaries were marked at the midpoint of the transition. The onset of the [r] was the most difficult to determine. It could often be recognised by the occurrence of noise, a change in period structure, or an increase in local energy at the boundary between the vowel and liquid (or a combination of these cues). If the change in period structure was gradual, the segment boundaries were drawn at the midpoint of the transition from vowel to liquid.

Recapitulated, the following three durational measurements were made:

- Alignment of H: the distance (in ms) of the peak from the end of the accented vowel of the test word.
- Alignment of L: the distance (in ms) of the valley preceding the peak (L) from the onset of the accented syllable of the test word.
- C0toC1: duration (in ms) from the onset of the accented syllable to the end of the accented vowel.<sup>4</sup>

This last measurement (duration of C0toC1) was made in order to see whether an effect of phonological length of the accented vowel on the

---

<sup>4</sup> Although the duration C0toC1 is not always equivalent to the duration of the accented syllable (see previous footnote), in the remainder of this chapter “C0toC1” and “duration of accented syllable” are used interchangeably.

alignment of the peak, could possibly be caused by a difference in the actual duration of the accented syllable. If one assumed that pitch movements are of a fixed duration (i.e. if the duration of a movement is constant), and the movement were to start at the onset of the accented syllable, then the endpoint of that movement would vary depending on the actual duration of that syllable. That is, if there is a fixed anchor point for the F0 rise at the beginning of the accented syllable, the occurrence of a short syllable forces the rise to peak later (assuming that the duration of the rise is constant) so that, by the time the F0 rise is completed, the speaker has already finished the short syllable and is into the following consonant. So, if the phonological differences between long and short syllables are also reflected in a phonetic difference (i.e. if the phonological long vowels are longer in duration than the phonologically short vowels), then the observed peak alignment differences could be caused by this phonetic difference in duration. Therefore, it was investigated whether there were differences in the actual duration of the accented syllable. Furthermore, additional analyses were carried out on a small subset of the materials in which this potential confound is absent (as will be explained further in the results section).

Figure 3.2 illustrates how the segments were marked and the various measurements were made.

## Design

The statistical package SPSS 8.0 for windows was used to run analyses of variance (ANOVAs). The statistical design was mixed (within and between items). There were two factors: LENGTH and SPEAKER. The factor LENGTH was between-items and had two levels (long and short). The factor SPEAKER was within-items and had five levels (D1, D2, D3, D4, and D5). All reported F-values are Greenhouse-Geisser epsilon corrected.

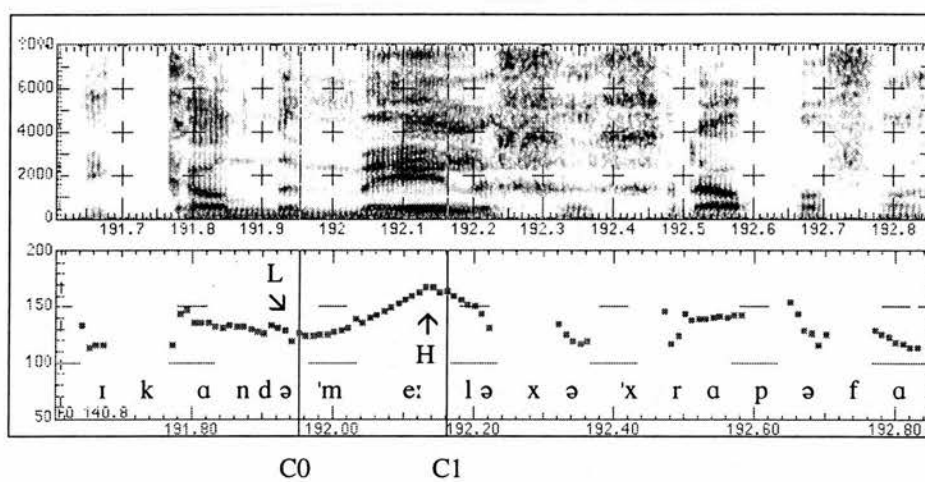


Figure 3.2. The various measurement points, illustrated with a sample sentence from experiment 1.

## 2.3. Results

### Alignment of H

To establish whether there was a difference in the alignment of H depending on the phonological length of accented vowel of the test word, the durational measurements for “alignment of H” for each speaker were entered into a two-way ANOVA (LENGTH X SPEAKER) with repeated measures on the variable SPEAKER. Table 3.4 presents the means and standard errors for the measurement “alignment of H” for each speaker (averaged over 20 items) in the different length conditions. The data are also graphed in Figure 3.3. The results of the ANOVA show that there was no significant main effect of SPEAKER ( $F < 1.5$ ). There was, however, a main effect of the factor LENGTH [ $F(1,35) = 42.139$ ;  $p < 0.0001$ ], with the H earlier (relative to C1) in the long condition (Set D-long) than in the short condition (Set D-short). There was no significant interaction between the two factors ( $F < 1.2$ ).

Table 3.4. Mean duration (ms) and standard error of “alignment of H” (relative to the end of the accented vowel) for each of the length conditions, averaged over 20 items per condition.

	A. Long condition		B. Short condition	
	mean	SE	mean	SE
Speaker D1	-11.33	7.48	29.58	7.28
Speaker D2	-16.78	4.45	18.79	4.33
Speaker D3	-1.28	7.20	27.53	7.01
Speaker D4	-21.22	5.50	25.79	5.35
Speaker D5	-9.28	6.89	22.05	6.71

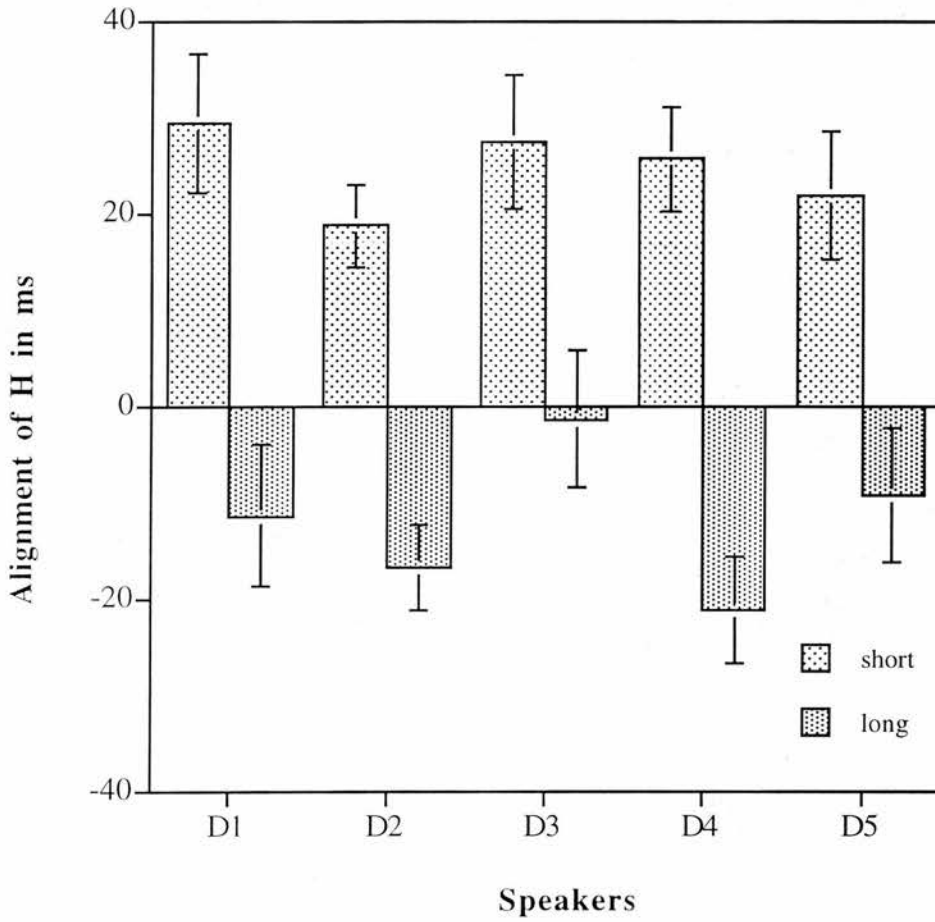


Figure 3.3. Mean duration (in ms) and standard error of "alignment of H" (relative to the end of the accented vowel) for each of the length conditions. Means are shown for each speaker of Group D.

## Alignment of L

Next, the alignment of the L was measured (i.e. the distance of the L from the onset of the accented syllable). A two-way ANOVA (LENGTH X SPEAKER) with repeated measures on the variable SPEAKER showed no significant effect of LENGTH [ $F < 1$ ]. However, it revealed a significant main effect of SPEAKER [ $F(1,4) = 3.610$ ;  $p < 0.011$ ], and a significant interaction between the two effects [ $F(1,4) = 2.711$ ;  $p < 0.039$ ].

Table 3.5. Mean duration (in ms) and standard error of "alignment of the L" (relative to the onset of the accented syllable) for each of the length conditions, averaged over 20 items.

	A. Long condition		B. Short condition	
	mean	SE	mean	SE
Speaker D1	1.17	5.29	-6.84	5.15
Speaker D2	11.06	6.64	-3.74	6.46
Speaker D3	-2.78	4.66	-15.84	4.54
Speaker D4	-3.00	4.74	10.16	4.61
Speaker D5	6.50	7.82	10.53	7.61

From Table 3.5 (which shows the means and standard error averaged over 20 items per condition) it can be seen that for two speakers (D4 and D5) the L is aligned later in the short than in the long condition. For the other three speakers, however, the L is earlier in the short than in the long condition. Separate one-way ANOVAs, however, showed that for none of the speakers this effect of LENGTH is significant (for each speaker [ $F < 2.7$ ]).

## Duration of accented syllable

Table 3.6. presents the mean duration of "C0toC1" for each speaker and each condition, averaged over 20 items. The duration measurement "C0toC1" was entered into a two-way ANOVA (SPEAKER X LENGTH)

with repeated measures on the factor SPEAKER. The results reveal that there is a significant effect of LENGTH [ $F(1,35) = 20.739; p < 0.0001$ ] as well as SPEAKER [ $F(1,4) = 18.519; p < 0.0001$ ], and no interaction between the factors [ $F < 1.3$ ]. Thus it appears that the phonological length of the accented vowel (long vs short) affected the duration of "C0toC1" (which in most cases is equivalent to the accented syllable). That is, the syllable duration is longer in the long condition (i.e. when the accented vowel is phonologically long) than in the short condition (i.e. when the accented vowel is phonologically short).

Table 3.6. Mean duration (ms) and standard error of "C0toC1" (in ms) for each of the length conditions, averaged over 20 items.

	Set A. Long condition		Set B. Short condition	
	mean	SE	mean	SE
Speaker D1	241.17	8.02	192.47	7.81
Speaker D2	217.28	9.59	156.11	9.34
Speaker D3	209.06	10.29	159.16	10.02
Speaker D4	206.06	8.47	154.58	8.25
Speaker D5	217.22	8.51	178.26	8.28

The significant effect of the duration "C0toC1" suggests that it is possible that the observed difference in alignment of H could in fact be a confound of the design, rather than an actual effect. In order to decide whether the effects observed could be explained by differences in the actual duration of the accented syllable, several analyses were carried out on a small subset of the data. The subset consisted of only the high long vowels [i, y] and their short counterparts [ɪ, ʏ], which are reported to have the same average phonetic duration (Booij, 1995: p. 5). The total number of sentences of this subset was 16 for each subject, i.e. 8 with the short and 8 with the long condition.

Table 3.7. Mean duration (ms) and standard error of "C0toC1" for each length condition, averaged over 8 items on the subset of data with the vowels [ɪ, ʏ, ɪ, and ʏ].

	Set A. Long condition		Set B. Short condition	
	mean	SE	mean	SE
Speaker D1	216.00	18.69	174.25	17.48
Speaker D2	187.29	15.39	149.25	14.39
Speaker D3	172.14	14.17	156.63	13.25
Speaker D4	188.29	17.16	143.75	16.05
Speaker D5	218.72	12.86	190.00	12.03

For the hypothesis (the H is aligned differently depending on the phonological length of the vowel) to be supported, the effect should still occur in this subset where there should be no difference in the segmental duration. To establish whether there are indeed no differences in the segmental durations across conditions, the durational measurements "C0toC1" were entered in a 2-way ANOVA (SPEAKER X LENGTH) with repeated measures on the factor SPEAKER. Table 3.7 shows the means and standard errors for each speaker and each condition (averaged over 8 sentences per condition). The results show that there is indeed no significant effect of LENGTH [ $F < 2.9$ ] on the duration of "C0toC1". There was, however, a significant effect of the factor SPEAKER [ $F(1,4) = 13.292$ ;  $p < 0.0001$ ], but no interaction between the factors [ $F < 1.3$ ].

Next, the measurement "alignment of H" was entered into a two-way ANOVA (SPEAKER X LENGTH). The means and standard errors are presented in Table 8, and graphed in Figure 3.4. The results reveal that there is a significant main effect of LENGTH [ $F(1,13) = 16.347$ ;  $p < 0.001$ ], with the H earlier in the long than in the short condition. The factor SPEAKER was not significant [ $F < 1.3$ ], and there was no significant interaction between factors [ $F < 2.2$ ].

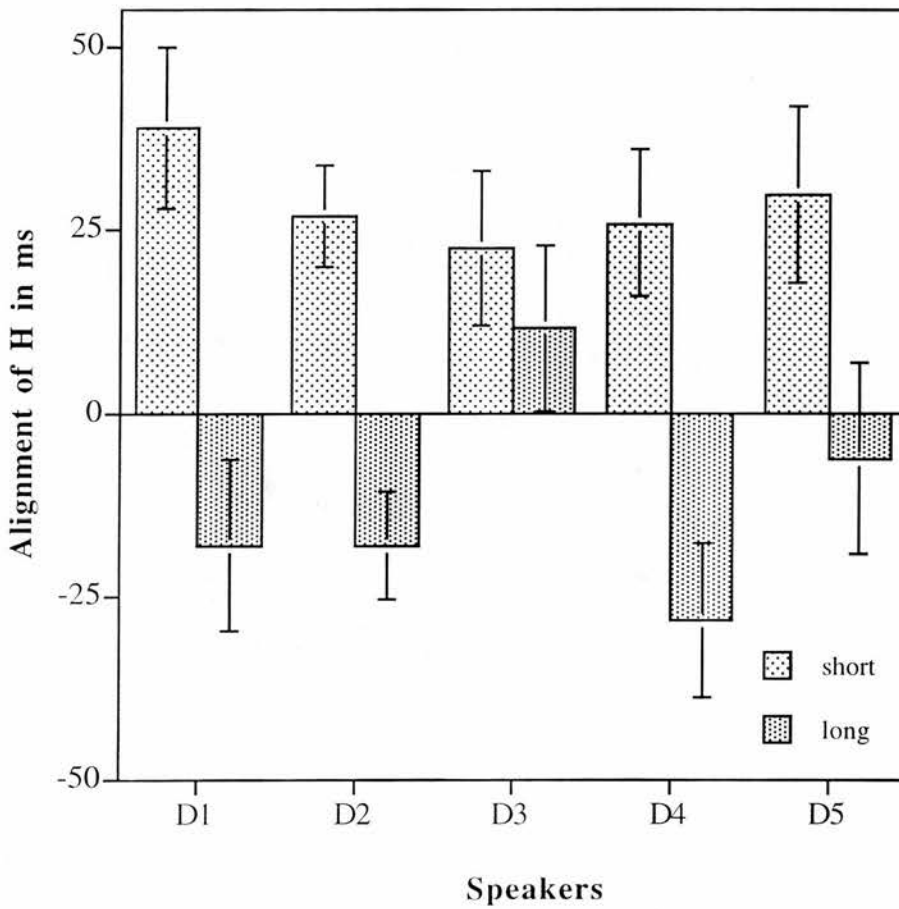


Figure 3.4. Mean duration (ms) and standard error of "alignment of H" for each of the length conditions. Means are shown for each speaker of Group D, and are based on a subset of the data with the vowels [ɪ, ʏ, ɪ, and ʏ].

Table 3.8. Mean duration (ms) and standard error of "alignment of H" for each length condition, averaged over 8 items on the subset of data with the vowels [ɪ, y, i, and ʏ].

	Set A. Long condition		Set B. Short condition	
	mean	SE	mean	SE
Speaker D1	-18.14	11.68	39.00	10.93
Speaker D2	-18.14	7.29	27.00	6.81
Speaker D3	11.57	11.40	22.63	10.67
Speaker D4	-28.29	10.70	26.00	10.01
Speaker D5	-6.14	13.04	29.88	12.20

#### 2.4. Summary of results

The main hypothesis tested in this experiment was that Dutch native speakers produce two different alignment patterns of the H in prenuclear accents, depending on the phonological length of the accented vowel of the test word. Strong support was found for this hypothesis, as in the main analysis there was a strongly significant difference between the alignment of H in the long and the short condition. There was also a strong effect of length on the duration of "C0toC1", suggesting that the difference in alignment of H could be due to a difference in segmental duration. However, in a subset of the data where the confound of a phonetic length difference was absent, the difference in alignment of H was still present. This finding suggests that it is the phonological length of the vowel which has an effect on the position of the peak in the syllable, rather than the actual phonetic length of that syllable. However, the examined dataset was rather small (8 sentences per condition, 5 speakers), and further research is therefore necessary to corroborate this finding.

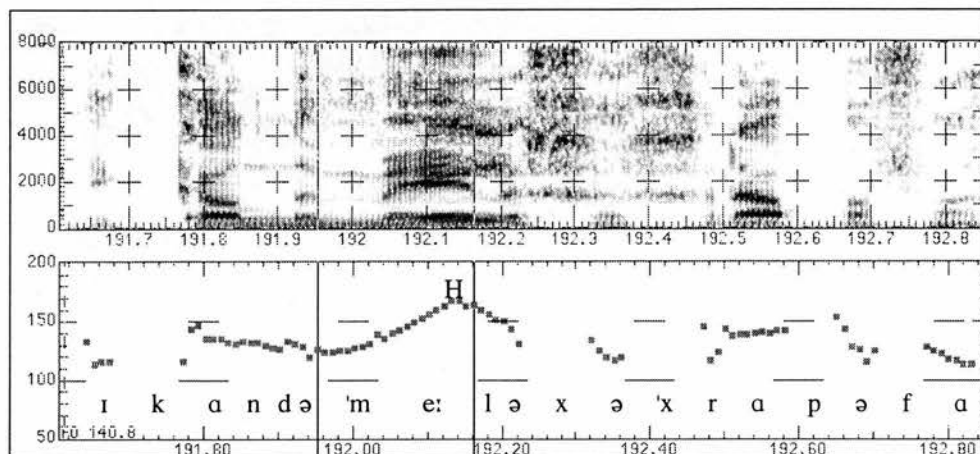
Since the aim of this experiment is to establish whether there are cross-linguistic differences in the alignment of the peak in Greek and Dutch prenuclear accents, the actual source of the observed alignment differences in Dutch is not our main concern. Therefore, this issue will

not be discussed in further detail here (but see Ladd, Mennen, and Schepman, forthcoming).

A second point examined in this experiment was whether there was also a difference in the alignment of the L across the two length conditions. There was no evidence in the data supporting that this was the case. It seems then that there is no effect of phonological vowel length on the position of the start of the rise in Dutch prenuclear accents.

Taken together, the results suggest that the starting point of the rise in Dutch prenuclear accents does not depend on the phonological length of the accented vowel, and seems to be located somewhere near the onset of the accented syllable. The endpoint of the rise is situated either in the accented vowel or in the following consonant, depending on the phonological length of the vowel. Figure 3.5 shows examples of the alignment in Dutch prenuclear accents, with an example of the long condition in panel (a), and an example of the short condition in panel (b). Finally, comparison of Figure 3.1 and 3.5 suggests that there is indeed a cross-linguistic difference in the alignment of H between Dutch and Greek prenuclear accents. One of the aims of the next experiment is to establish whether these small differences are indeed significant.

(a)



(b)

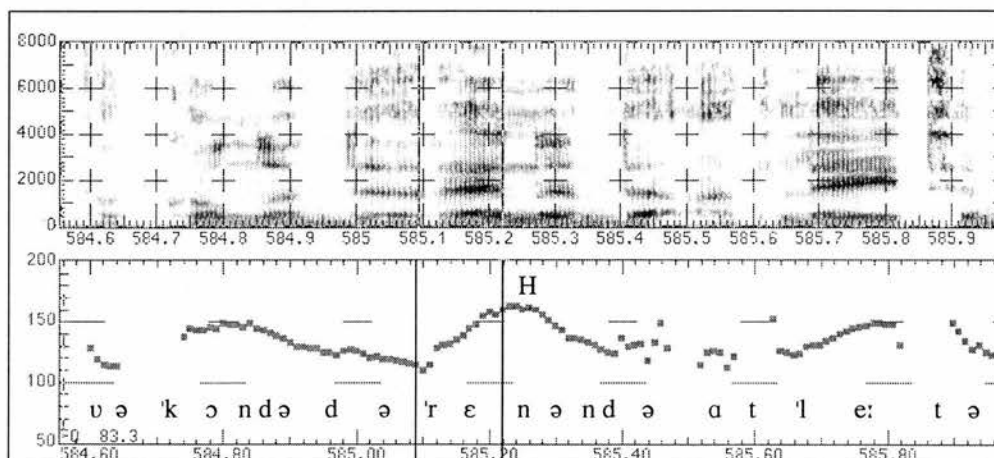


Figure 3.5. Spectrogram and F0 trace of an example of (a) the long condition, and (b) the short condition. The sentences shown are from experiment 1 (no. 26 and 5 in appendix A respectively). The peak is located in the accented vowel in the long condition, but in the following consonant in the short condition. The vertical lines delimit the stressed syllable of the word bearing the prenuclear accent.

### 3. Experiment 2

#### 3.1. Introduction

In experiment 1 it was suggested that there are cross-linguistic differences in the alignment of the peak in Greek and Dutch prenuclear accents. The existence of such cross-linguistic differences raises the question as to whether such differences can be acquired by second language learners. In this experiment, the L2 alignment of non-native (Dutch) speakers of Greek will be compared against native Greek speakers' alignment in Greek and native Dutch speakers' alignment in Dutch. In this way it will be established whether L2 learners, when acquiring the intonational system of an L2, develop a 'merged' system. A 'merged' system, where L2 learners merge the properties of the L1 and the L2 sounds, has been observed in the L2 acquisition of segments (e.g. Flege and Hillenbrand, 1984). The notion of a 'merged' system has been posed by Flege & Hillenbrand (1984) to explain segmental aspects of L2 acquisition. It is not known from the existing literature whether it could explain L2 acquisition at the prosodic level.

The following research questions are addressed in this experiment.

1. Can L2 learners (in this case Dutch speakers of Greek) ever acquire native-like values of Greek peak alignment?
2. If L2 learners do not fully acquire Greek peak alignment, do they develop a 'merged' system, with values intermediate between native L1 and L2 values?
3. Is it possible to explain the data of the L2 learners by Flege's Speech Learning Model (SLM)?

The question whether L2 learners will ever acquire native-like values of L2 peak alignment, is related to the issue of language experience and ultimate attainment. Many studies have suggested that increased

exposure to the second language may improve the production and perception of L2 speech (Best and Strange, 1992; Bohn and Flege, 1990; Flege, 1987; Flege, 1995; Oyama, 1976; Sheldon and Strange, 1982; Strange, 1995). However, language exposure is not a guarantee for accent-free L2 pronunciation (Scovel, 1969). Some researchers assume that adult L2 learners will never be able to speak a second language without a trace of foreign accent. In order to come to such a conclusion, it is crucial to look at the data of L2 learners who have reached the most advanced stages of interlanguage development. Only then will we be able to decide on the degree of success attainable by adult second language learners. The group of non-native (Dutch) speakers of Greek (Group DG) which participated in this study are all adult L2 learners who have reached an 'ultimate level of attainment', since they had many years (an average of 20.8 years) of language experience.<sup>5</sup>

The first and second question are related to the degree of 'phonetic similarity' between the two languages under consideration (Flege, 1995). According to Flege (1995) learners cannot develop a new perceptual target when the sounds in L1 and L2 are 'similar', and place the L2 sound in the same category as the similar L1 sound. This will result in inaccurate production. When, on the other hand, the sound is noticeably different from any phonetic category in the L1, learners will develop a new category for this 'new' sound, and accurate production should be possible.<sup>6</sup> Flege's model has been developed to account for segmental aspects of L2 learning only. In the present study, we will investigate whether his model can also account for differences at the suprasegmental level (question three). As was seen before, both Dutch and Greek prenuclear accents can be seen as instances of a H\* accent. However, they differ in phonetic detail, with the

---

<sup>5</sup> See further Chapter 1 (section 3.4) for a discussion on the issue of ultimate attainment and near-nativeness.

<sup>6</sup> There is, however, no guarantee that the L2 learner will actually achieve accurate production of a 'new' L2 sound. It is possible that, even when a new category is established, the L2 sounds differs when compared to the native norm. It is thought that in order to maintain contrast within and across languages, the categories may deflect away from each other (see Flege, 1995: p. 242, Flege et al, 1995: p.3133; and chapter 1 of this thesis).

H\* occurring earlier in Dutch than in Greek. In Flege's terminology the Greek prenuclear accent is 'similar' to the Dutch prenuclear accent. As a result, category formation for the L2 alignment will be blocked and a single category will be used for both the L1 and L2 peak alignment.

If Flege's model is correct and if it can be applied to the suprasegmental level (question 3), we would expect that even advanced L2 learners would not achieve native-like values for this aspect of L2 intonation. Instead, alignment would be inaccurate, and intermediate in degree between the L1 and L2 norm ('merged' system). That is, the peak of the non-native speakers of Greek should be earlier than that of native speakers of Greek, but later than that of native speakers of Dutch.

### 3.2. Method

#### Materials

The items consisted of three sets of 20 sentences, two of which (Set D-long, and Set D-short) were identical to those used in Experiment 1, and one (Set Greek) which was originally designed and analysed for a previous study on the alignment of (native) Greek prenuclear accents (cf. Arvaniti, Ladd, and Mennen, 1998). The three sets were:

- Set D-long: a set of 20 Dutch sentences with phonologically long vowels [i, e, a, o, y] in the accented syllable of the test word, in equal distribution.
- Set D-short: a set of 20 Dutch sentences with phonologically short vowels [ɪ, ε, ɑ, ɔ, ʏ] in the accented syllable of the test word, in equal distribution. The vowels used in this set were paired with the vowels used in set D-long. The long-short vowel pairs of Set D-long and D-short are /i, ɪ/, /e, ε/, /a, ɑ/, /o, ɔ/, and /y, ʏ/.
- Set Greek: a set of 20 Greek sentences constructed with the full set of five vowels in the accented syllable of the test word in roughly

equal distribution. This set of materials is a subset of the 25 test sentences designed for Arvaniti et al's experiment 2 (1998).

Unlike Dutch, which has 16 vowels with both phonologically long and short vowels as well as diphthongs, Greek has a simple five-vowel system, consisting of /i, e, a, o, u/, all of which are of equal phonological weight (e.g. Joseph and Philippaki-Warburton, 1987; Mirambel, 1959).

Table 3.9. *Experiment 2: sample test items for Set Greek; the test words are underlined (adapted from Arvaniti, Ladd, and Mennen (1998)).*

1.	[to <u>poðilato</u> tu 'petru 'klapice xtes 'vraði 'ekso a'po to 'spiti tu] (The bike Petros's stolen3sg yesterday evening from his house) "Petros's <u>bike</u> was stolen last night outside his house."
2.	[ta <u>yliko'lemona</u> pu a'yorases ðen 'itan 'freska] (the limes that bought2sg. not was fresh) "The <u>limes</u> you bought were not fresh."
3.	[i pa'remvasi tu ipur'yu den 'efere to poði'to apo'telezma] (The intervention minister's not brought3sg the desired result) "The minister's <u>intervention</u> did not have the desired result."

As noted in the method section of Experiment 1, the Dutch items were designed to match the Greek set as closely as possible, in order to allow comparisons between the sets. Therefore, the long vowels of the vowel pairs used in the Dutch sets were in most cases the same as those used in the Greek set. However, since in Dutch there is no obvious corresponding phonologically short vowel to the vowel /u/, the /u/ was not used in the Dutch sets but another vowel pair was chosen (i.e. /y, ʏ/). As noted in Experiment 1, in all items the test words were exclusively words with lexical stress on the antepenultimate syllable and there were two to five unaccented syllables following the stressed syllable. As described in Experiment 1, the syllable structure was also matched, and for ease of measurement only sonorants and voiced fricatives were used in the relevant syllables of the test words. Examples of the items included in the

Greek set are given in Table 3.9 (examples of the two Dutch sets were already shown in Table 3.3). A full list of the Greek test items can be found in Appendix B, a full list of the Dutch test items is given in Appendix A.

### Subjects

Three groups of subjects were used in this experiment, (i) five native speakers of Dutch (Group D), (ii) five native speakers of Greek (Group G), and (iii) five non-native (Dutch) speakers of Greek (Group DG).

- Group D: the same native speakers of Dutch who participated in Experiment 1 also took part in Experiment 2. They were three female (D1, D4 and D5) and two male subjects (D2 and D3).
- Group G: Two males (G1, G2), and three females (G3, G4, and G5), all recruited from the Edinburgh student population, with no self reported hearing and reading problems, participated in this experiment. They were all in their twenties, and at the time of recording had been in Edinburgh for periods ranging from a few months to four years. All of them were educated at university level, and were brought up in Athens. They all spoke Greek with a standard Athenian accent.
- Group DG: The five non-native (Dutch) speakers of Greek were three males (DG1, DG3 and DG5), and two females (DG2 and DG4). (The results of a third female speaker (DG6) had to be discarded, because she produced disfluencies to such an extent that it was not possible to obtain a sufficient number of test items.) The speakers were all native speakers of Dutch and were very experienced speakers of Greek, with between 12 and 35 years of language experience. For reasons of confidentiality the age and amount of experience with the L2 of each individual speaker will not be revealed here. They had all started learning Greek after the age of puberty. All five speakers held a university degree in Modern

Greek Language and Literature, and currently teach Greek at university level.

As mentioned in Experiment 1, it was an unrealistic criterion for speaker selection to use only monolingual speakers for Group D and Group G. Therefore, all speakers of all three groups were also reasonably competent in English. None of the speakers was a near-native speaker of another language, except for the speakers of Group DG, who were all near-native speakers of Greek. All speakers were selected by the present author (who is a native speaker of Dutch, and also is a near-native speaker of Greek).

### Procedure

In general the experimental procedure was similar to that described in Experiment 1. There were, however, small differences in procedure between the three groups of speakers. For reasons of clarity, the procedure for Group D (which has already been described for Experiment 1) is repeated here:

- Group D: The native speakers of Dutch read each test sentence (of Set D-long and Set D-short) twice from a randomised set of cards, each card containing one typed sentence. In order to avoid listing effects, the cards were interspersed with cards which contained material for other (pilot) experiments, which will not be reported here. The recording session took approximately 15 minutes.
- Group G: The native speakers of Greek read each test sentence (Set Greek) twice from a randomised list typed in Greek. Practice sentences were added at the beginning (seven fillers) and the end (five fillers) of the list, to avoid discourse-initial and final F0 effects. The materials recorded in this session were exclusively materials from set Greek, no other materials (for other experiments) were

interspersed. The entire recording session lasted approximately 15 minutes.

- Group DG: For this experiment the non-native (Dutch) speakers of Greek were recorded in their L2 Greek. The recordings were supervised by the author who interacted with the speakers in Greek. Before the materials were recorded the author held a short conversation with the speakers in Greek. The speakers were then presented with a set of cards, each card containing one sentence typed in Greek. The speakers read each test sentence of the Greek set (Set Greek) twice from the randomised set of cards. In order to avoid listing effects, the cards were interspersed with cards which contained material for experiments 4, and 5 (see chapter 4 for a detailed description of these experiments), and for a further experiment which will not be reported here. Consequently the sentences of each of the four experiments acted as fillers for the other. The recording of the Greek items lasted approximately 40 minutes. When the speakers had finished reading the Greek materials, they were asked to take a short break (about ten to fifteen minutes). The author immediately switched to the Dutch language, and continued addressing the speakers in Dutch. After the break, the speakers continued with the recording session, but now materials for experiment 3 were recorded (which will be reported in the method section of experiment 3, below).

### Apparatus and measurements

All materials were recorded on professional equipment at different locations: in the studio of the Department of Phonetics, University of Amsterdam; in the studio of the Department of Linguistics, University of Groningen; in the recording studio of the Department of Linguistics, University of Edinburgh; or at a quiet room in the speaker's home. As some of the speakers of Group DG requested that their identity should not

be revealed, it is not revealed here which speakers were recorded at which location. None of the speakers was paid for their participation.

The recordings for all the speakers were made on digital audio tape (DAT). Data selection, digitisation, and measurements were done as described in Experiment 1. The following measurements were made:

- Alignment of H: the distance (in ms) of the peak from the end of the accented vowel of the test word.
- Alignment of L: the distance (in ms) of the valley preceding the peak (L) from the onset of the accented syllable of the test word.

## Design

There was one factor (GROUP), with three levels (D, DG, and G). In order to keep the design identical for the different comparisons, the design was between items (even though Group DG and G produced the same set of items), since the items differed from those produced by Group D.

### 3.3. Results

The data were analysed in the following way: First, for each group the means were calculated for each item of the different sets. These means were then entered into two one-way ANOVAs: one comparing the group means for Group DG (Greek) with those of Group D (D-long), and one comparing them with the means of Group D (D-short). Next, in another one-way ANOVA the means of Group DG (Greek) were compared with those of Group G (Greek).

Since one of the research questions was to establish whether L2 learners could achieve native-like alignment in the L2, it was necessary to

not only analyse the group means, but also analyse the data of each speaker separately. This was done by separate ANOVAs, which compared the data of each speaker of Group DG with the group means of Group G.

### 3.3.1. Group analyses

#### Alignment of H

The means and standard errors of the measurement “alignment of H” for each Group are presented in Figure 3.6. A one-way ANOVA comparing Group D (D-long) with Group DG (Greek), showed that there was a significant effect of the factor GROUP, with the peak earlier in Group D than in Group DG [ $F(1,38) = 4.880; p < 0.033$ ]. However, there was no significant effect of GROUP when the short condition (D-short) was compared with the means of Group DG (Greek) [ $F < 2.1$ ].

However, a one-way ANOVA comparing the means of Group DG (Greek) with those of Group G (Greek), revealed that there was a significant effect of the factor GROUP, with the peak later in Group G than in Group DG [ $F(1,38) = 80.503; p < 0.0001$ ].

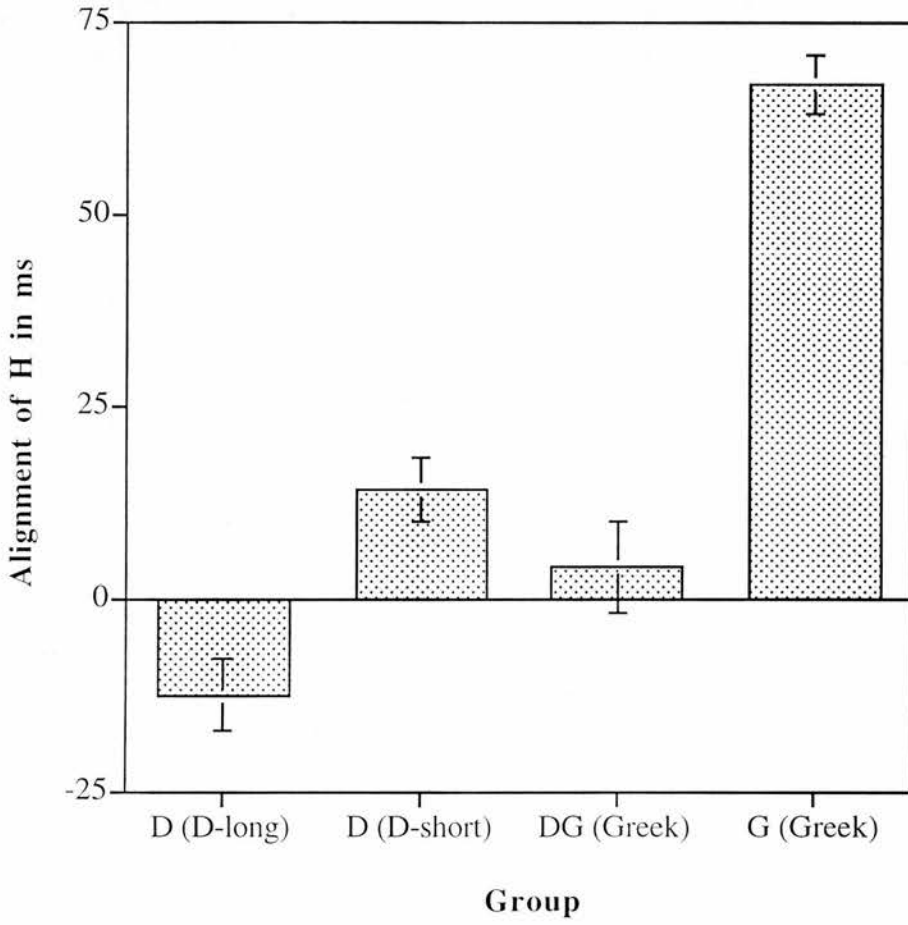


Figure 3.6. Mean duration (ms) of "alignment of H" (relative to the end of the accented vowel) for Group D (D-long), Group D (D-short), Group DG (Greek) and Group G (Greek).

## Alignment of L

The same analyses as described above for “alignment of H” were conducted, but now with “alignment of L” as the dependent variable. There was no significant main effect for the factor GROUP in any of the analyses: for Group D (D-long) vs Group DG (Greek) [ $F < 1$ ]; for Group D (D-short) vs Group DG (Greek) [ $F < 2.4$ ]; and for Group DG (Greek) vs Group G (Greek) [ $F < 1.64$ ].

### 3.3.2. Individual peak alignment analyses

Table 3.10. Mean duration (ms) and standard error of “alignment of H”, for each speaker of Group DG, together with the group means of Group G.

	mean	SE
Speaker DG1	-3.84	12.97
Speaker DG2	-20.28	9.65
Speaker DG3	-28.95	6.75
Speaker DG4	58.30	8.07
Speaker DG5	6.60	9.70
Group G	67.12	3.94

To establish the amount of success of each individual speaker on achieving L2 alignment values, data for each speaker were subjected to a one-way ANOVA, comparing their individual data (on set Greek) with the group means for Group G (on the mean Group values for set Greek). Since there was no significant difference between the group means of Group DG and Group G for the alignment of L, only the alignment of H was subjected to individual analyses. The results from individual ANOVAs show that there is a main effect of the factor GROUP for speakers DG1 [ $F(1,37) = 39.804; p < 0.0001$ ], DG2 [ $F(1,36) = 75.749; p < 0.0001$ ], DG3 [ $F(1,38) = 151.214; p < 0.0001$ ], and DG5 [ $F(1,33) = 40.430; p < 0.0001$ ], but not for speaker DG4 [ $F < 1$ ]. The means and standard errors for

each subject are presented in Table 3.10, and graphed in Figure 3.7. As can be seen from table 3.10 and Figure 3.7, the peak alignment values for speaker DG4 lie within the norms for the native speakers of Greek. The values for the other speakers, however, are significantly different from those of Group G, i.e. their peak alignment is significantly earlier than that of Group G and speaker DG4.

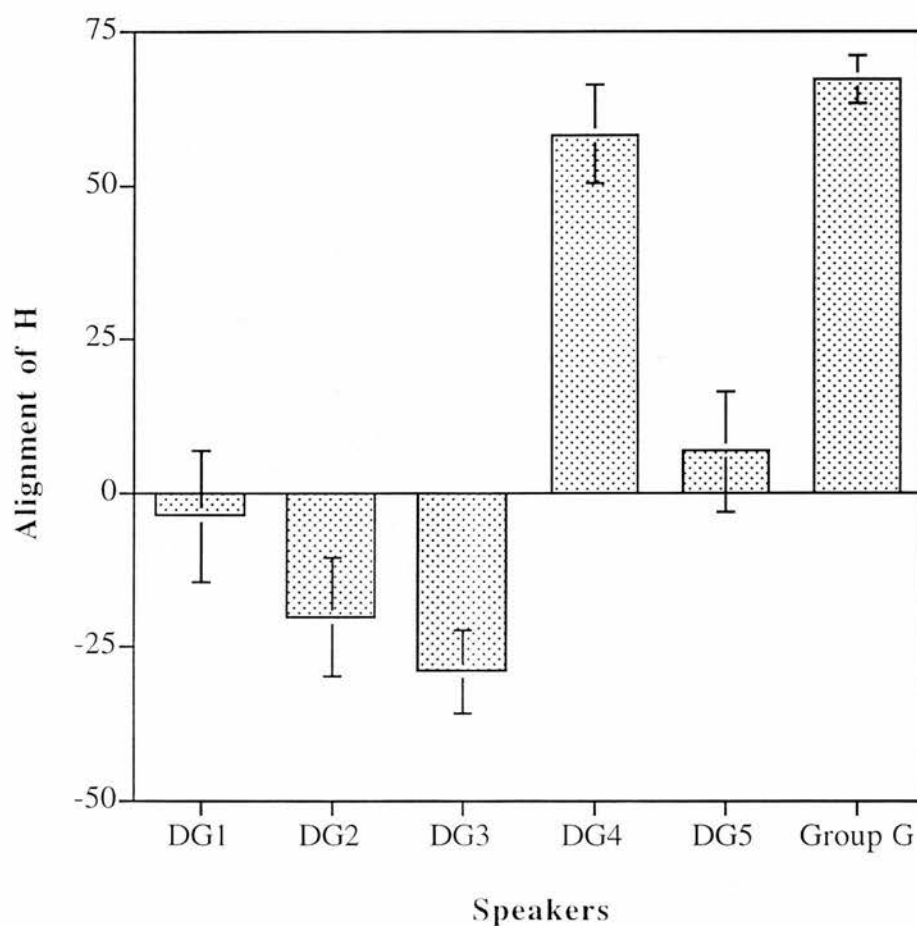


Figure 3.7. Mean duration (ms) of "alignment of H" (relative to the end of the accented vowel) for each speaker of Group DG (Set Greek), together with the group means for Group G (Set Greek).

### 3.3.3. Additional group analyses

In section 3.3.1. it was seen that Group DG's peak alignment was significantly earlier than that of Group G, and that it was significantly later than the long condition (D-long) of Group D, but not significantly different from the short condition (D-short) of Group D. However, in the individual speakers' analyses it was shown that speaker DG4 was different from the other speakers of Group DG, and actually achieved native Greek peak alignment values. Therefore, her data may have obscured the actual tendency of the other speakers of Group DG. Therefore, the analysis reported in section 3.1.1 is redone here, but now the means for group DG are based on four speakers only (i.e. the data for speaker DG4 are left out of the analysis). An ANOVA comparing Group D (D-long) with Group DG (Greek), reveals that now there is no effect of the factor GROUP ( $F < 0.01$ ). However, there is a significant effect of GROUP when Group D (D-short) is compared to Group DG (Greek) [ $F(1,38) = 12.750; p < 0.001$ ]. There is also still an effect of GROUP when Group DG (Greek) is compared to Group G (Greek) [ $F(1,38) = 121.065; p < 0.0001$ ]. In other words, when the data of speaker DG4 are left out, for Group DG the peak is aligned significantly earlier than that of Group G, and significantly earlier than the short condition (D-short) of Group D. The alignment of the peak is not significantly different from that of the long condition (D-long) of Group D. The means and standard errors for the different groups and conditions are graphed in Figure 3.8.

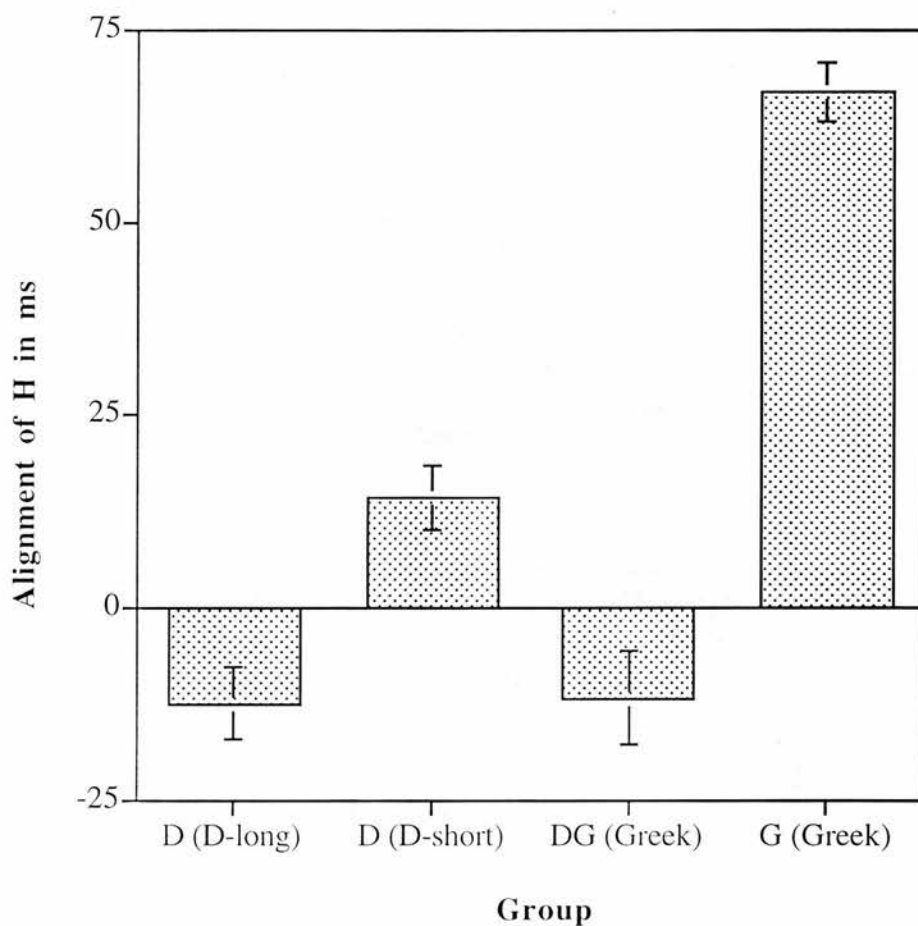


Figure 3.8. Mean duration (ms) of "alignment of H" (relative to the end of the accented vowel) for Group D (D-long), Group D (D-short), Group DG (Greek) and Group G (Greek). The means for Group DG are calculated from the data of speakers DG1, DG2, DG3, and DG5 (DG4 is excluded).

### 3.4. Summary of results and discussion

In this experiment, the L2 alignment of non-native (Dutch) speakers of Greek (Group DG) was compared to native Greek speakers' alignment in Greek (Group G) and native Dutch speakers' alignment in Dutch (Group D). The aim of this comparison was to establish whether L2 learners, when acquiring the intonational system of an L2, develop a 'merged' system.

Alignment measurements were taken for both the peak (alignment of H, relative to the end of the accented vowel of the test word) and the preceding valley (alignment of L, relative to the onset of the accented syllable of the test word).

It was found that there were no significant overall differences in the alignment of the L between the different groups. For all the groups and all conditions the L seemed to be located somewhere near the onset of the accented syllable. Previous studies have found a similar consistency in the alignment of the F0 movements with the segmental string. Specifically, a number of recent papers have demonstrated that the F0 minimum at the beginning of a rising pitch accent consistently aligned with the onset of the accented syllable (e.g. Caspers and Van Heuven, 1993 for Dutch, Prieto et al., 1995 for Mexican Spanish, Arvaniti et al., 1998 for Greek).

There were, however, significant differences in the alignment of H. The overall analyses of peak alignment showed significant differences between the means of Group DG and Group G, suggesting that Group DG had not acquired native Greek values of peak alignment. However, the individual analyses showed that one non-native speaker (DG4), contrary to expectation, managed to produce peak alignment values which were within the norm for native speakers of Greek. Although Flege's model predicts that production of 'similar' L2 sounds will be inaccurate even after long exposure to the L2, it does not preclude full mastery of L2 sounds (Flege, 1995). However, speaker DG4 was by no means the most

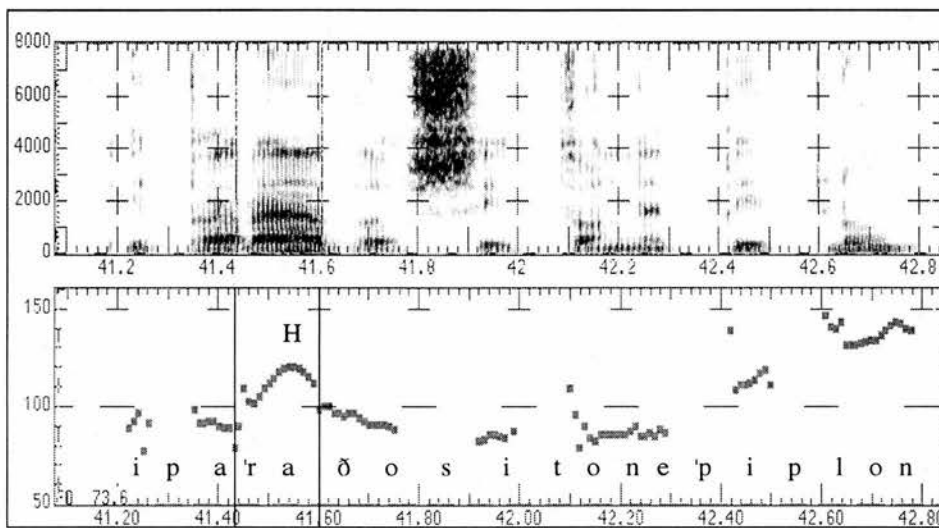


Figure 3.9. Spectrogram and F0 trace of the Greek prenuclear accent [i pa'raðosi ton e'piplon] "The delivery of the furniture" (sentence 2, Appendix B), spoken by one of the non-native speakers (DG3). The peak (H) is earlier than that produced by Greek native speakers (as illustrated in Figure 3.1). The vertical lines delimit the stressed syllable of the word bearing the prenuclear accent.

experienced L2 speaker, and it is not clear why she was able to achieve native levels, whereas some of the more experienced speakers were not as successful.

On the basis of the SLM it was predicted that L2 learners would develop a 'merged' system, that is they would produce peak alignment values in the L2 which would be later than those for Group DG and earlier than those for Group D. There was no evidence in the data supporting that this was the case. The means for the less successful speakers (DG1, DG2, DG3, and DG5) were not different from those obtained from the long condition (D-long) of Group D. This seems to suggest that Group DG has simply transferred its Dutch peak alignment (when the accented vowel of the word bearing the prenuclear accent is long) to the L2. However, in order to verify this assumption, the L2 data of the speakers of Group DG need to be compared to their L1 peak alignment data. In Figure 3.9 an example is shown of early alignment in a Greek prenuclear accent by a speaker of Group DG.

#### 4. Experiment 3

##### 4.1. Introduction

This experiment was designed to test whether experience with the L2 can have an effect on the L1, specifically on its peak alignment. According to the SLM (e.g. Flege, 1992), L2 learners are expected to classify 'similar' L1 and L2 sounds into a single category, due to the mechanism of equivalence classification (see Chapter 1, section 3.1). If applied to our intonation data, the SLM would predict that non-native (Dutch) speakers of L2 Greek, will place the 'similar' prenuclear accents into a single category. The use of a single category for both Dutch and Greek prenuclear accents would result in inaccurate production of the L2 prenuclear accent. The results of Experiment 2 showed that for most of the L2 speakers the production of the similar Greek prenuclear accent was indeed inaccurate.

However, the SLM also predicts that the production of a 'similar' L1 sound will gradually shift away from the monolingual norm. In other words, not only will L2 learners fail to achieve the native-language norm in the L2, their L1 production will also be affected.

It is assumed that this influence of the L1 on the L2 and vice versa, is caused by the fact that L1 and L2 phonological systems are not fully isolated (e.g. Flege, 1995: p. 241; Major, 1990: p.15-16; see also Chapter 1, section 3.4). As a result of this bidirectional interference, it is predicted that the better L2 learners become at approaching (in our case) native-like peak alignment, the more their peak alignment in the L1 will be affected. The aim of this experiment is to test whether this prediction is supported in the data.

A further aim of this experiment is to discover what the nature of this bidirectional influence or transfer is. In experiment 2 it was found that most of the speakers of Group DG did not acquire native Greek peak alignment values. However, there was also no evidence for a 'merged' system, intermediate between native Dutch and native Greek peak alignment values. So if speakers do not achieve native-like values, nor seem to develop a merged system, does this indicate that they simply transfer their L1 alignment to the L2? And if this is true, which of the two alignment patterns of the L1 is transferred and why? Furthermore, if we were to find evidence of bidirectional influence (i.e. L1 influencing L2, and L2 influencing L1) how would this be reflected in the interlanguage of Group DG?

## 4.2. Method

### Materials

The items were 40 Dutch sentences with prenuclear accents on the test word, identical to those used in experiment 1. They were of two types:

D-long: 20 sentences with phonologically long vowels in the accented syllable of the test word.

D-short: 20 sentences with phonologically short vowels in the accented syllable of the test word.

For more detailed information about the materials used the reader is referred to the method section of Experiment 1.

## Subjects

The materials were recorded by two groups of subjects, group D (native speakers of Dutch) who had taken part in experiment 1, and group DG (Dutch non-native speakers of Greek), who had taken part in experiment 2. For a more detailed description of the groups the reader is referred to the method sections of experiment 1 and 2.

## Procedure

- The procedure for Group D is described in Experiment 1.
- Group DG: the non-native (Dutch) speakers of Greek recorded the two sets of Dutch test sentences (D-long and D-short) in the same session as the set of Greek sentences (Set Greek) were recorded. The Greek sentences were recorded first, followed by a short break (about ten to fifteen minutes). The procedure for the Greek set is described in the method section of Experiment 2. During the break the author who supervised the recording session chatted to the speakers in Dutch. After the break, the Dutch test sentences were recorded. The experimental procedures for the Dutch test sentences were exactly the same as for the first group of speakers (native

speakers of Dutch), described in Experiment 1. The recording of the Dutch set lasted approximately 15 minutes.

### Apparatus and measurements

The apparatus and measurements were identical to that described in experiment 1, with the exception that now only the alignment of H was measured (relative to the end of the accented vowel).

### Design

The design was mixed (between and within items). There were two factors: GROUP and LENGTH. The factor GROUP was within items and had two levels (D and DG). The factor LENGTH was between items, and also had two levels (long, and short). Since we were also interested in the amount of L2 influence on the L1 for each subject, in addition to the general analyses, separate ANOVAs were performed for each speaker of Group D. All reported F-values are Greenhouse-Geisser epsilon corrected.

### 4.3. Results

First, for each group of speakers the mean values of "alignment of H" were calculated for each item and entered into a two-way ANOVA (GROUP X LENGTH), with repeated measures on the factor GROUP. The results show that there is no significant effect of GROUP ( $F(1,38) < 1$ ). There was, however, a significant effect of LENGTH [ $F(1,38) = 36.020$ ;  $p < 0.0001$ ], together with a significant interaction between the two factors [ $F(1,38) = 10.519$ ;  $p < 0.002$ ]. The means and standard errors are graphed in Figure 3.10. As can be seen in Figure 3.10, this interaction is due to the fact that there is a larger difference between the two conditions for Group D

than for group DG. That is, in the short condition alignment is earlier for Group DG than for Group D, whereas in the long condition it is later for Group DG. In other words, for Group DG the alignment in the two conditions are closer together than for Group D.

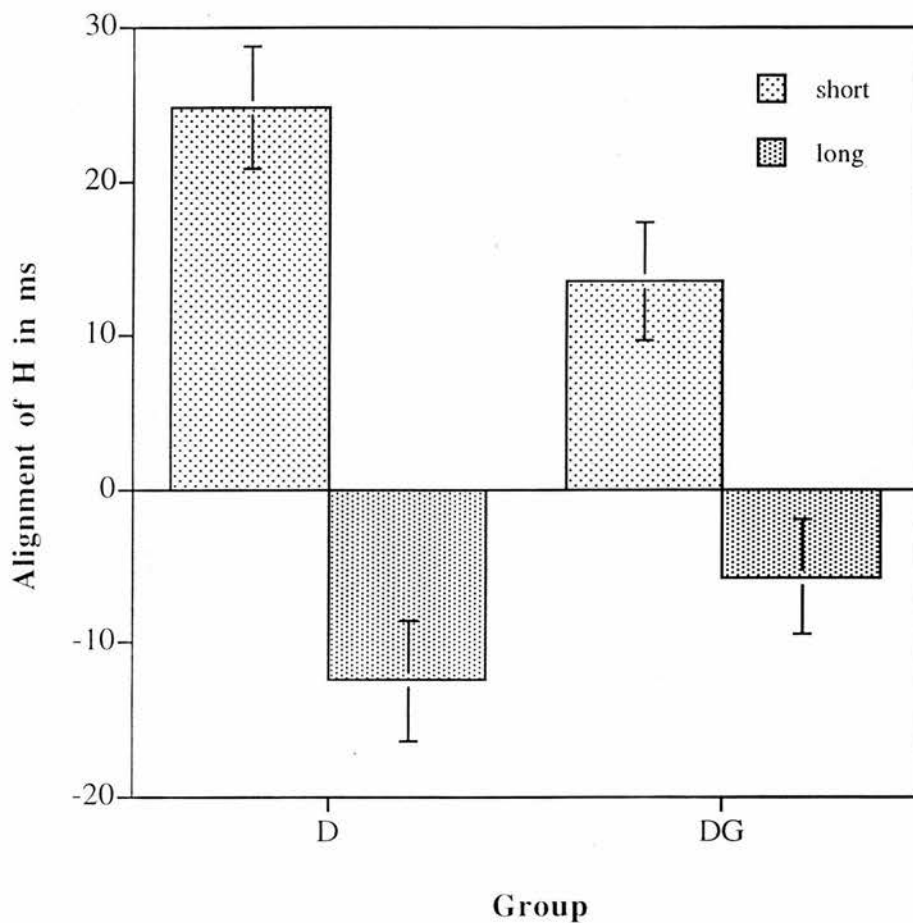


Figure 3.10. Mean duration (in ms) and standard error of "alignment of H" (relative to the end of the accented vowel) for Group D and Group DG in the long and short condition, averaged over 20 items.

The results of separate ANOVAs for each speaker of Group DG, confirm that the speakers of Group DG do not show exactly the same pattern as Group D. The speakers of Group D all show a different alignment pattern in the long and short condition (as described in the result section of experiment 1), with the H earlier in the late than in the short condition. However, only speakers DG1 and DG4 of Group DG show a similar effect of LENGTH (for DG1 [ $F(1,38) = 5.621$ ;  $p < 0.023$ ], and for DG4 [ $F(1,37) = 16.292$ ;  $p < 0.0001$ ]). For the other speakers of Group DG there is no significant effect of LENGTH (for DG2 [ $F(1,36) = 4.033$ ;  $p = 0.052$  ns], for DG3 [ $F(1,38) = 2.963$ ;  $p = 0.093$  ns], and for DG5 [ $F(1,36) = 0.717$ ;  $p = 0.40$  ns]). Table 3.10 presents the means and standard errors for all speakers of each Group. The data are also graphed in Figure 3.11.

Table 3.10. Mean duration (ms) and standard deviation for the “alignment of H” (relative to the end of the accented vowel), for each speaker of Group D and Group DG..

	A. Long condition		B. Short condition	
	mean	SE	mean	SE
Speaker D1	-11.33	7.48	29.58	7.28
Speaker D2	-16.78	4.45	18.79	4.33
Speaker D3	-1.28	7.20	27.53	7.01
Speaker D4	-21.22	5.50	25.79	5.35
Speaker D5	-9.28	6.89	22.05	6.71
Speaker DG1	-6.15	6.96	18.85	7.92
Speaker DG2	-26.89	6.86	-8.11	6.36
Speaker DG3	-9.15	3.81	.70	4.27
Speaker DG4	-5.20	7.06	33.95	6.61
Speaker DG5	17.79	7.54	23.68	6.12

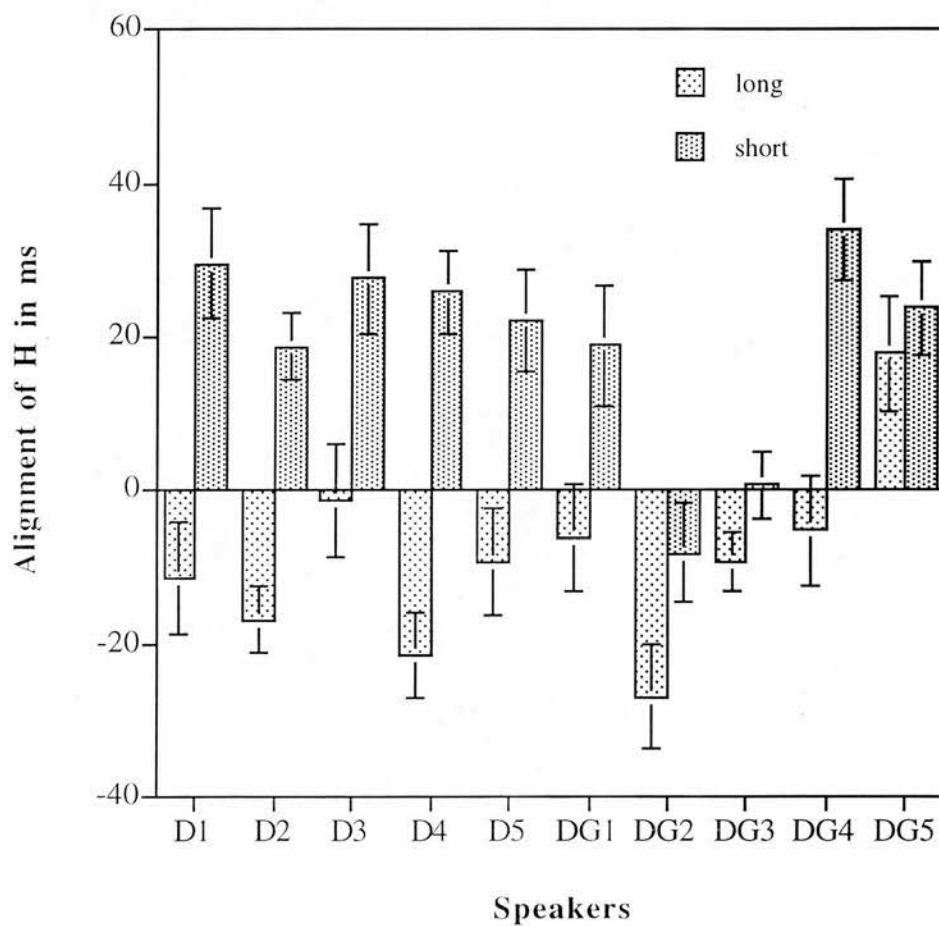


Figure 3.11. Mean duration (in ms) of "alignment of H" (relative to the end of the accented vowel) for each of the speakers of Group D and Group DG in the long and short condition, averaged over 20 items.

## Transfer

In order to see whether there is transfer from the L1 system to the L2 of the speakers of Group DG, the peak alignment in their L1 data (Dutch) was compared to that in their L2 (Greek) data. For each speaker two one-way ANOVAs were run, one comparing the Dutch long condition (Set D-long) to the Greek condition (Set Greek), and the other comparing the Dutch short condition (Set D-short) to the Greek condition. In this way, it could be determined which of the two Dutch conditions, if any, was transferred to Greek. The first analysis (D-long vs. Greek) showed that for all speakers except DG4 there is no significant difference between the Greek peak alignment and the alignment in the Dutch long condition (for DG1 [ $F(1,37) < 1$ ], for DG2 [ $F(1,36) < 1.25$ ], for DG3 [ $F(1,37) < 1$ ], and for DG5 [ $F(1,32) < 1$ ]). However, the second analysis (D-short vs. Greek) showed that for three of the speakers there was also no significant difference between the Greek peak alignment and the Dutch peak alignment in the short condition (for speaker DG1 [ $F(1,37) < 3$ ], for speaker DG2 [ $F(1,35) < 1.2$ ], for speaker DG5 [ $F(1,32) < 2.5$ ]). Speaker DG3 showed a significant difference between the Greek peak alignment and the Dutch peak alignment in the short ( $[F(1,38) = 13.783; p < 0.001]$ ), but not in the long condition (see above), suggesting that he transferred the alignment of the Dutch long condition to Greek. That is, for Speaker DG3 the peak is earlier in Greek than in the D-short condition, but there is no difference between alignment in Greek and in D-long. Speaker DG4's Greek peak alignment data are significantly different from both the short ( $[F(1,37) = 5.385; p < 0.026]$ ) and long ( $[F(1,38) = 35.052; p < 0.0001]$ ) Dutch conditions. This is as expected, as speaker DG4 has managed to acquire native Greek peak alignment values, and is therefore not expected to show transfer from the L1 to the L2. Her peak alignment values show that the peak is earliest in D-long, later in D-short and the latest in Greek. The means and standard errors for the individual speakers are presented in Table 3.11 and graphed in Figure 3.12.

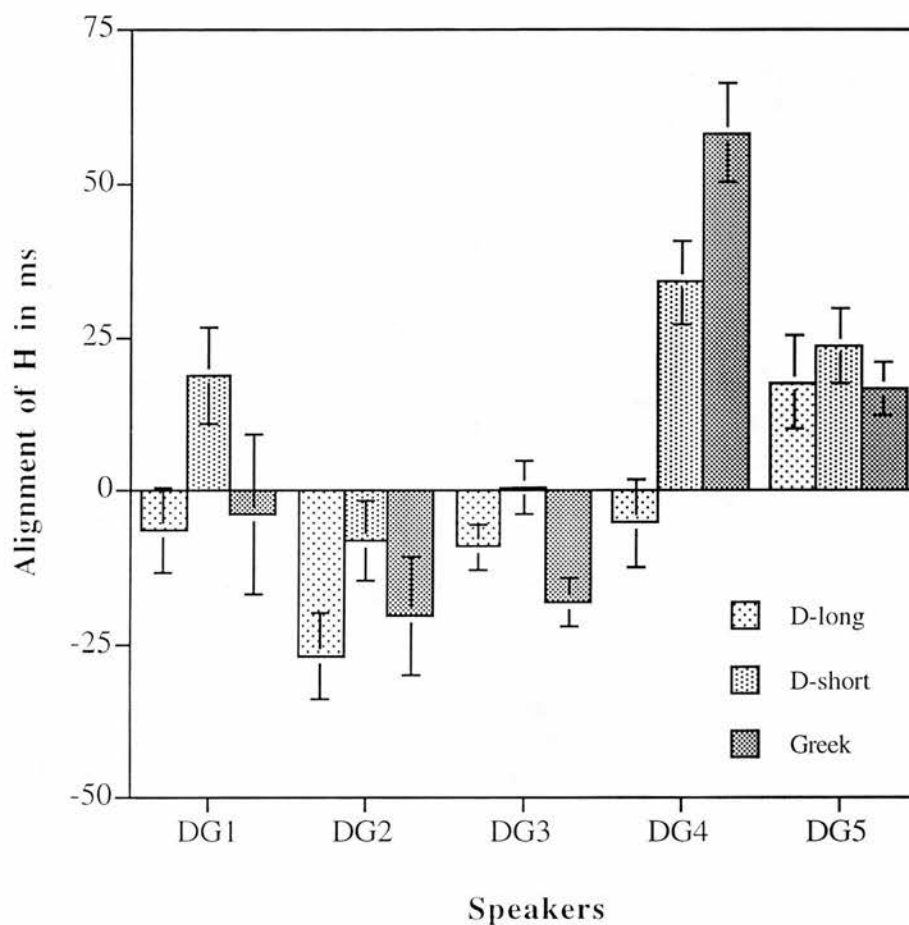


Figure 3.12. Mean duration (ms) of "alignment of H" (relative to the end of the accented vowel) for each of the sets of test items, D-long, D-short and Greek. The means are given for each of the speakers of Group DG.

Table 3.11. Mean duration (ms) and standard error for the “alignment of H” (with respect to the end of the accented vowel), for each speaker of Group DG in the three conditions: D-long, D-short, and Greek..

	D-long		D-short		Greek	
	mean	SE	mean	SE	mean	SE
Speaker DG1	-6.15	6.96	18.85	7.92	-3.84	10.77
Speaker DG2	-9.15	3.81	-8.11	6.36	-12.32	3.94
Speaker DG3	-26.89	6.86	.70	4.27	-18.24	3.87
Speaker DG4	-5.20	7.06	33.95	6.61	58.30	5.40
Speaker DG5	17.79	7.54	23.68	6.12	16.74	4.45

To summarise, the results reveal that the speakers who have failed to acquire native Greek patterns of alignment (all speakers except speaker DG4) differ in the way they align the peak across conditions. Their alignment pattern is either (i) not significantly different across the three conditions (DG1, DG2, DG5), or (ii) there is a significant difference in alignment between the D-short and Greek conditions, but no difference between D-long and Greek (DG3). Speaker DG4 is altogether different as her data show a significant difference across all three conditions, i.e. she has acquired native-like peak alignment in the L2 and at the same time has maintained a contrast between the long and short condition in the L1.

#### 4.4. Summary of results and discussion

This experiment tested whether experience with the L2 can affect peak alignment in the L1. For this reason, the alignment of the peak in Dutch prenuclear sentences for Group D (who had no extensive experience with an L2) was compared to that of Group DG (who had between 12 and 35 years of experience with L2 Greek). For each group the alignment of H was measured in two conditions, (i) test sentences with a long vowel in the

accented syllable of the test word (D-long), and (ii) test sentences with a short vowel in the accented syllable (D-short).

It was found that both groups showed a significant difference in alignment between the long and the short condition. However, there was also a significant interaction between the factors group and length, indicating that for Group DG the peak in the short conditions is earlier, whereas in the long condition it is later than that of Group D. That is, for Group DG the difference between the alignment in the long and the short condition is smaller than that of Group D. Furthermore, individual speaker's analyses showed that of the speakers of Group DG who had failed to acquire native-like values of peak alignment in the L2, only one speaker (DG1) showed a difference between the long and short condition in his L1. For none of the other speakers (who had not acquired native-like values of peak alignment in the L2) such a distinction was found. This is rather different from the results of Group D, where all speakers aligned the peak significantly earlier in the long than in the short condition. These results suggest that the L1 of the group of non-native (Dutch) speakers of Greek is indeed affected.

The main hypothesis tested was the one put forward by Flege (1995) and Major (1990), who argue that L1 and L2 phonological systems are not fully isolated and can therefore influence one another. They suggest that as a result of this it is likely that the better the L2 learners become at approaching the L2 norm, the more their L1 will be affected. For this hypothesis to be true, speaker DG4 (who actually achieved native values of peak alignment) should show the most effect on her L1. However, the results of speaker DG4 show that although she has mastered L2 peak alignment, she is still able to maintain the distinction between alignment in the long and the short condition in her L1. The other speakers of Group DG, on the other hand, have not approached native values of L2 peak alignment, yet their L1 alignment pattern is different from that of native L1 speakers who do not have extensive knowledge of an L2.

Although this experiment shows that the L1 peak alignment of Group DG is different from that of Group D, we cannot be entirely sure that this difference is caused by Group DG's experience with the L2. Although unlikely, the possibility remains that this difference already existed before L2 learning started.

A second aim of this experiment was to test what the nature of the observed bidirectional influence was. It was hypothesised that if speakers failed to reach native-like norms in the L2, but also did not develop a merged system, it was likely that they had transferred L1 values to the L2. However, which of the two L1 alignment categories were transferred to the L2? Furthermore, if there was also an effect of the L2 on the L1, how would this be reflected in their interlanguage system?

In order to test this the data of Group DG on all three conditions (D-long, D-short and Greek) were compared. In the individual analyses there appeared to be no significant difference between the Dutch long condition and the Greek condition, suggesting that speakers transfer the D-long alignment to Greek. This supports the SLM's hypothesis that L2 learners tend to use just one category for similar L1 and L2 sounds (in this case contours).

However, the situation seems to be slightly more complicated than that. This experiment showed that some speakers also do not distinguish between the peak alignment in D-short and Greek. In other words, for these speakers there is no difference in peak alignment between D-long, D-short and Greek. This suggests that some speakers use just one alignment category for all three conditions.

## 5. Rate and duration

In experiment 1 it was briefly mentioned that a possible explanation for the difference in peak alignment in Dutch prenuclear accents was the actual duration of the accented syllable. It was argued that if one assumed

that the risetime were constant (i.e. if the duration of a pitch movement were of a fixed duration), then the endpoint of the rise (H) should be placed earlier when the duration of the syllable is longer and later when it is shorter. However, analyses on a subset of data, where the duration of the accented syllable was equal amongst the long and short condition, showed that there was still an effect of phonological length of the vowel on the alignment of H. Therefore, it is unlikely that this effect is due to a difference in duration.

In a similar vein it could be argued that the difference in peak alignment between native (Group G) and non-native speakers of Greek (Group DG) could be explained by a difference in speaking rate or duration. If the risetime were equal for both groups, the H should be placed earlier for the group that speaks more slowly, than for the group that speaks faster. A number of studies on L2 acquisition have suggested that an L2 is often produced more slowly than is the native language (e.g. Lennon, 1990; Munro and Derwing, 1995; Pennington, 1992; Raupach, 1980; Sabin *et al.*, 1979). It is therefore likely that the non-native speakers of this study were slower in Greek than were the native Greek speakers.

In order to test this some additional measurements were made for the two groups (DG and G) in the Greek set only. The measurements were:

- LtoH : the distance (in ms) from the start of the rise (L) to the end of the rise (H), i.e. the risetime.
- C0toC1: duration (in ms) from the onset of the accented syllable to the end of the accented vowel (which is in most cases also the end of the accented syllable).

First for each group the means were calculated for each item. Then for each measurement the means were entered into a one-way ANOVA (repeated measures). The results for the dependent variable C0toC1 show that there is indeed a significant difference between the two groups, with Group DG speaking more slowly than Group G [ $F(1,19) = 74.275$ ;  $p < 0.0001$ ]. However, the results for the dependent variable LtoH show that

there is also a difference in risetime between the two groups: the risetime is shorter for Group DG than for Group G [ $F(1,19) = 11.727$ ;  $p < 0.003$ ]. The means and standard errors for both measurements and both groups are presented in Table 3.12.

Table 3.12. Mean duration (ms) and standard error for the two measurement "C0toC1" and "LtoH" for Group DG and Group G.

	Group DG		Group G	
	mean	SE	mean	SE
C0toC1	195.65	8.46	148.34	6.09
LtoH	199.66	5.38	220.51	7.45

To summarise, Group DG speaks slower but has a faster risetime than Group G. The fact that the risetime of Group DG is faster than that of Group G makes it less plausible that the peak alignment differences found are caused by the difference in rate between the two groups. The results show that the assumption that there is a "fixed" risetime (i.e. a risetime which is of a certain invariant duration) in the present data can be rejected. In fact, Arvaniti *et al.* (1998) already put forward that it is unlikely that pitch movements are of a fixed duration. Their results indicate that:

"...the timing and scaling of the beginning and ending of the prenuclear accentual rise in Greek is not determined by properties of the rise qua pitch movement, but rather the exact opposite is true. The L and the H of the accent are anchored to segmentally defined positions, and the duration and slope of the pitch movement are completely determined by the segmental composition of the accented word." (p.24)

In any case, it can be concluded that the differences in segmental durations between Group G and Group DG, cannot fully explain the differences in alignment which were found.

## 6. General discussion

The series of experiments reported in this chapter was aimed at empirically testing the production of prenuclear Greek accents in declaratives by non-native (Dutch) speakers of Greek. The first experiment reported in this chapter was aimed at providing evidence for the observation that the alignment of the peak in Dutch prenuclear accents is affected by the phonological length of the lexically stressed vowel of the word bearing the prenuclear accent. The results of this experiment show that this is indeed the case the H is earlier (near the end of the accented vowel) when the accented vowel is phonologically long, and later (in the following consonant) when that vowel is phonologically short. It was suggested that it is unlikely that this difference in alignment was due to a difference in the actual duration of the accented syllable, as an analysis of a subset of the data (where there was no difference in the duration of the accented syllable) still yielded the same results.

As previous research had suggested that in Greek prenuclear accents the peak is usually situated in the postaccentual vowel (Arvaniti *et al.*, 1998), the results of this experiment confirmed that there are cross-linguistic differences in peak alignment between Greek and Dutch prenuclear accents.

In experiment 2 it was investigated whether Dutch advanced speakers of L2 Greek have acquired native-like values of Greek peak alignment, and if not, whether they have developed a merged system with values intermediate between the L1 and L2. The results show that even though the speakers of Group DG had many years of experience with the L2, most of them had failed to acquire native-like values for peak alignment. This finding appears to support the position put forward in Flege's SLM (e.g. 1995), which claims that if L1 and L2 sounds (in this case contours) are 'similar', the L2 sound will be placed in the same category as the similar L1 sound, which will result in inaccurate production. According to the SLM, language exposure is no guarantee for accurate production of L2

segments. The results of this experiment suggest that this position may be upheld for L2 prosody, at least for L2 peak alignment.

However, experiment 2 also showed that one speaker managed to acquire peak alignment values which were within the range of values produced by native speakers of Greek. In other words, she had acquired native values of peak alignment. Although the SLM suggests that it is unlikely that L2 learners will acquire accurate production of a 'similar' L2 sound, it does not exclude the possibility that some speakers achieve native-like values. Specifically, the SLM posits that "L2 sounds will eventually be produced as specified in phonetic category representations. If the new phonetic category established by a bilingual for an L2 sound matches native speakers', then the L2 sound will be produced accurately" (Flege, 1995: p. 240). However, it does not specify whether this is only true for "new" sounds, or also for "similar" sounds. Furthermore, it does not explain why some speakers manage to achieve accurate production of an L2 sound, and others do not, even though they may have had the same amount (or less) of language experience.

It is tempting to speculate as to what may have caused this speaker to be more successful than the others. On the basis of the data now in hand her success cannot be explained. However, it is worth pointing out that an informal interview revealed that she started learning the L2 by means of naturalistic exposure in Greece (language immersion), without any formal instruction. Furthermore, she was slightly younger (18 years old) than the other speakers (who started learning Greek from the age of 20 to 25) when she started learning Greek. By contrast, the other non-native speakers had all started out with formal instruction in the L2 (in the Netherlands), and initially did not stay in Greece. It is possible that DG4's success can be attributed to this difference in learning situation, or to the slightly earlier age at which L2 learning began. In fact, there are some studies which indicate that full mastery of L2 pronunciation and intonation is attainable in adulthood (Neufeld, 1977; Neufeld, 1987). There are also many studies that suggest that the earlier in life one begins

to learn a L2, the better one is apt to pronounce it (e.g. Flege, Munro, and Mackay, 1995; Long, 1990; Oyama, 1976). Furthermore, these studies suggest that the critical period for speech does not end abruptly, but that there is a linear relation between non-native's age of L2 learning and their degree of perceived foreign accent in that L2. It may be that the difference in the age of L2 learning between DG4 and the other speakers of Group DG can explain her success. However, since the difference in age of L2 learning is only two years whereas the difference in the degree of success achieved is rather large, this explanation seems rather unlikely.

However, on the basis of these limited data no definite claims about the effect of age or of the learning situation can be made, but further research into this possibility may be worth pursuing.

The next question which was addressed in this study was whether L2 peak alignment reflects some kind of intermediate system, with peak alignment values somewhere between the L1 and L2 norm. It was argued that if Group DG had developed a merged system, their peak would be earlier than that observed for Group G, but later than that of Group D. This prediction was not supported by the data. Peak alignment for Group DG was as early as that of Group D in the Dutch long condition (i.e. when the accented vowel of the test word was long). It appears that the SLM's position that L2 learners develop a merged system (Flege and Hillenbrand, 1984; Flege, 1995) cannot be supported by the present data on L2 prosody.

In experiment 3 it was investigated whether experience with an L2 may have an effect on the peak alignment in the L1. It was shown that for all of the speakers of Group D there was a significant effect of phonological length on the peak alignment. However, although the same effect was found for Group DG, it was shown that this difference in alignment was smaller than that of Group D. Furthermore, it was shown that this alignment difference did not show up in the data of all the speakers of Group DG. In fact, it was shown that most speakers of Group DG did not show this pattern. These results support the SLM's (e.g. Flege,

1995) and Major's (1990) position that interference is bidirectional and that L2 learning may have an effect on the L1.

For speaker DG4, however, no evidence was found to support such bidirectional interference, since speaker DG4 showed native values of peak alignment in *both* the L1 and the L2. The SLM assumes that influence of the L2 on the L1 is more obvious in advanced L2 learners, and argues that the better L2 learners become at approaching native-like values, the more their L1 will be affected. If this were true, speaker DG4 should show the greatest effect on the L1. Based on this finding, it is possible to reject this position.

Another position is put forward by Major (1990), who claims that L2 learners can either (i) fail to achieve accent-free L2 speech and maintain native L1 pronunciation; (ii) achieve native-like L2 pronunciation but lose native L1 pronunciation; or (iii) lose native L1 pronunciation but still fail to achieve native-like L2 pronunciation. The results for most of the speakers of Group DG are compatible with the latter possibility. However, the results of speaker DG4 are not compatible with any of the possibilities suggested by Major, as she shows evidence of nativeness in *both* the L1 and the L2 (at least for this temporal aspect of L2 intonation).

The last question which was addressed was how this bidirectional interference was reflected in the interlanguage system of Group DG. It was shown that the speakers who have failed to acquire native-like peak alignment values (all speakers except speaker DG4) have L2 peak alignment values that are similar to their alignment values in the Dutch long condition. This suggests that they transfer the L1 peak alignment of the long condition (D-long) to the L2. However, for most speakers L2 peak alignment was also not significantly different from their alignment values produced in the Dutch short condition, although their L2 values were closer to the D-long than to the D-short condition. Only speaker DG4 showed a significant difference in alignment between all three conditions (D-long, D-short, and Greek), whereas most of the other speakers use just one alignment category for all three conditions. It is not clear how this

could be explained by the SLM. The SLM predicts that one category will be used for both L1 and L2 sounds which are perceived as similar. For our data it would predict that one category would be used for both L1 and L2 peak alignment. However, if speakers were to classify the Greek peak alignment into the D-long category, nothing would prevent them to still maintain a contrast between long and short peak alignment in their L1 (as for example was the case for speaker DG1). However, most speakers use just one category for all three contrasts.

In the next chapter, further experiments are reported where the production of intonation by non-native (Dutch) speakers of Greek is compared to that of native Greek speakers. In these experiments the production of another sentence type (i.e. yes/no questions), will be studied.

## Chapter 4

### Nucleus placement in L2 Greek yes/no questions

#### 1. Introduction

In the previous chapter it was found that most non-native (Dutch) advanced speakers of (Modern) Greek had failed to acquire native peak alignment in Greek prenuclear accents. The two experiments described in this chapter are intended to provide instrumental data to test the proficiency of non-native speakers of Greek in their production of Greek yes/no questions.

In Chapter 3 it was shown that there are cross-linguistic differences between yes/no question intonation in Dutch and Greek. To recapitulate, the following differences between Dutch and Greek yes/no question intonation can be found:

- Dutch has several yes/no question patterns,  $H^* (L^-) H\%$  or  $L^* (H^-) H\%$ , whereas Greek has only one pattern,  $L^* H^- L\%$ .
- the nuclear accent in Greek yes/no question intonation is  $L^*$ , in Dutch it can either be  $L^*$  or  $H^*$ .
- the peak of the rise-fall in Greek yes/no questions occurs *after* the nuclear accent, whereas in Dutch it is always associated with the nuclear or prenuclear accent.
- in Greek yes/no question intonation the  $H^-$  phrase accent is affected by the location of the nucleus: if the nuclear accent is on the last word, the  $H^-$  occurs on the sentence's *final syllable*; if it is on an earlier word, the  $H^-$  occurs on the *stressed syllable of the last*

*word*. There is no obvious counterpart of the H<sup>-</sup> phrase accent alignment in Dutch.

- in Greek yes/no questions the boundary tone is usually L%. In Dutch yes/no questions the boundary tone is usually H%.
- the neutral location for the nuclear accent in Greek yes/no questions is the verb, whereas in Dutch having the main accent on the verb would make the question non-neutral.

In Chapter 2 it was also concluded that it is not that straightforward to apply the notions of 'new' and 'similar' posited in Flege's (1992) Speech Learning Model (SLM) to differences in Dutch and Greek yes/no question intonation. It was argued that in terms of phonetic similarity (on the basis of which the SLM determines the relation between L1 and L2 sounds) it has to be concluded that the perceptual distance between Dutch and Greek yes/no questions may not be equally great for each focus reading (NF or NNF). It was argued that the nucleus-final (NF) Greek yes/no questions should be seen as 'new', whereas the nucleus-non-final (NNF) yes/no question should be seen as 'similar'. This argument was based on the fact that in Greek NNF yes/no questions the peak of the final rise-fall (i.e. the H<sup>-</sup> phrase accent) always occurs on the lexically stressed syllable of the utterance-final word. Since in Dutch a rise-fall always associates with a lexically stressed syllable, this pattern is similar to Dutch. A rise-fall on an unstressed syllable, on the other hand, is not possible in Dutch. Therefore, it was argued that the NF pattern should be seen as new.

In this chapter, two experiments are reported. They test the production of Greek yes/no question intonation by non-native (Dutch) speakers of Greek. The non-native speakers are all very advanced speakers of Greek, and have all started learning the L2 after puberty. The first experiment (experiment 4) tested their production on the 'similar' NNF yes/no questions. The second (experiment 5) tested the production of the 'new' NF yes/no questions, and compared the results with those obtained for the NNF yes/no questions.

## 2. Experiment 4: nucleus non-final yes/no questions

### 2.1. Introduction

The aim of this experiment is to examine the production of Greek nucleus-non-final (NNF) yes/no questions by non-native (Dutch) speakers of Greek. As noted above, according to the SLM (Flege, 1992), Greek NNF yes/no questions should probably be considered 'similar' to Dutch. The SLM posits that when L1 and L2 sounds are acoustically similar (but not identical), both sounds will be classified into a single category, which will result in inaccurate production of the L2 sound. Although the SLM was developed to account for segmental aspects of language learning, the results of the experiments described in chapter 3 suggest that it may be useful to explain certain prosodic aspects of L2 learning.

In this experiment the applicability of the SLM to L2 intonation will be further explored. The SLM would predict that the group of non-native (Dutch) advanced speakers of Greek would fail to recognise the differences between the final rise-fall in Greek NNF yes/no questions and a Dutch rise-fall. In Dutch a rise-fall is always associated with a stressed syllable, just as in Greek NNF yes/no questions. However, it is likely that Dutch L2 speakers would fail to recognise that the H peak in Greek NNF yes/no questions is not a nuclear or prenuclear H\* accent, but rather a phrase accent which in fact occurs *after* the nuclear accent. It is therefore predicted that even advanced L2 speakers will fail to produce this L2 pattern accurately.

## 2.2. Method

### Materials

A corpus of 60 sentences with Greek yes/no questions was used, in which the nuclear accent was expected to be placed on the final content word of the sentence.<sup>1</sup> The sentences were short and contained a maximum of two content words. In addition, the position of the lexical stress of the final content word was varied. This manipulation of lexical stress was done in order to test the hypothesis that the placement of the rise-fall is affected by the position of the last stressed syllable in NNF, which was borne out by the data of the native speakers of Greek (Arvaniti, Ladd, and Mennen, forthcoming), but had not been tested for non-native speakers. Thus, the test sentences were divided into three sets of 20 sentences each; one in which lexical stress was on the utterance-final syllable, a second in which lexical stress was on the penultimate syllable, and a third in which it was on the antepenultimate syllable of the final content word.<sup>2</sup>

In order to achieve the desired nucleus placement, sentences were presented in short dialogues. An example of such a dialogue is given below:

[na sas si'stiso] "Let me introduce you to each other."

[ti ynɔrɪzɪs ti ma'rina] "Have you met Marina?"

The speakers were not given any explicit instructions to achieve the desired nucleus placement. Under these conditions, in 99% of items the

---

<sup>1</sup> The materials described in experiment 4 and 5 have previously been used in another study, which investigated native Greek yes/no question intonation. The findings on aspects of native Greek yes/no question intonation are reported in Arvaniti, Ladd, and Mennen (forthcoming).

native speakers of Greek placed the accent where it was expected. As mentioned in Chapter 2, the neutral location for the nuclear accent in Greek yes/no questions is the verb, whereas in Dutch having the main accent on the verb would make a question non-neutral. It was therefore expected that non-native speakers would experience difficulty in producing the desired nucleus placement, and may not place the accent where it was expected. This will be discussed further in the results section. It was not expected that the L2 speakers would experience difficulties with the correct placement of *lexical* stress, as in Greek orthography the stress is indicated on the appropriate syllable.

Table 4.1. *Experiment 4: sample test items for nucleus-non-final (NNF) yes/no questions with different stress patterns of the final content word (adapted from Arvaniti et al (forthcoming)). The test word is underlined.*

Nucleus location	Final word stress pattern	Sample test item	Gloss
NNF	final	[θa fas <u>karame'le</u> ]	"Will you have some <u>crème caramel</u> ?"
	penultimate	[tu arese to 'ðoro]	"Did he like the <u>present</u> ?"
	ante-penultimate	[to θimase to 'numero]	"Do you remember the <u>number</u> ?"

All sentences had a minimum of two unstressed syllables separating the stressed syllable of the final word from the preceding stressed syllable. This was done in order to avoid the effect of stress clashes on the realisation of pitch accents as those found in Arvaniti et al. (1998). The test words contained mostly sonorants in the relevant syllables (i.e. the stressed syllable and following consonants of the final content word), so

<sup>2</sup> In Greek, lexical stress is allowed on any one of the last three syllables of a word, but no further to the left (Joseph and Philippaki-Warbuton, 1987; Malikouti-Drachman and Drachman, 1980; Mirambel, 1959).

that the F0 contour would be uninterrupted and it would be fairly easy to determine the location of F0 minima and maxima. Examples of some of the test sentences are given in Table 4.1. A full list of items can be found in Appendix C.

## Subjects

Two groups of subjects were used, a group of native Greek speakers (Group G) and a group of non-native (Dutch) speakers of Greek (Group DG):

- Group G consists of five native speakers of Greek, three females (G3, G4 and G7) and two males (G6, G10).<sup>3</sup> Two of the female speakers (G3, G4), had also taken part in the experiments described in Chapter 3 (Experiments 2 and 3). All speakers were recruited from the Edinburgh student population and had been resident in Edinburgh for periods ranging from a few months to four years. They all spoke Greek with a standard Athenian accent, except for speaker G7 who had been brought up in the Peloponnese and had a slightly different accent. It was thought that her accent did not affect her intonational system, but see later for a more detailed analysis. All speakers could communicate well in English, but none of them could be considered near-native (see chapter 1, section 3.4 for a discussion on the issue of ultimate attainment and near-nativeness).
- Group DG consists of the same speakers who participated in Experiment 2 and 3 (Chapter 3). This time, however, we did not

---

<sup>3</sup>Speakers G3, G4, G6, G7, and G10 of the present study, are referred to as KA, DA, KP, VP and TV, respectively in (Arvaniti, Ladd, and Mennen, forthcoming). The data of a sixth speaker (AA) described in that study were discarded as she was considered to be a near-native speaker of English, and this could possibly have influenced her intonation pattern in the L1 (see further Chapter 3).

have to exclude the third female speaker (DG6, whose data we had to discard in the previous experiments) as in this experiment she produced a sufficient amount of data that could be analysed. This brought the number of speakers of Group DG to six, three males (DG1, DG3, and DG5) and three females (DG2, DG4, and DG6). All speakers were very advanced speakers of Greek (with an average of 17 years of experience with the L2), and had all started learning Greek after puberty.

None of the speakers of the two groups had any speech or hearing impairment and they were all naive as to the purpose of the experiment.

### Procedure

Each test sentence was presented on a separate card, and was incorporated in a very short dialogue (as described above). The cards were presented in blocks of twenty dialogues in random order. For Group DG the cards with the dialogues were interspersed with cards which contained materials for experiment 2 (see Chapter 3), experiment 5 (see below) and a pilot experiment which is not reported here. Therefore, the materials of each experiment acted as fillers for the other. For Group G, the cards were interspersed with materials from another experiment, described in Arvaniti et al. (forthcoming), which will not be discussed in this thesis. The entire recording session took approximately 40 minutes for Group DG, for Group G it lasted approximately 50 minutes.

## Apparatus and measurements

All items were recorded on digital audio tape (DAT) on professional equipment in the recording studio of the Department of Linguistics, University of Edinburgh (Group G); and in the recording studio of the Department of Phonetics, University of Amsterdam or in a sound-attenuated booth of the Department of Linguistics, University of Groningen (Group DG). For reasons of privacy it is not revealed here which speakers were recorded in Groningen and which in Amsterdam.

Recordings and digitisation were identical to those described in Experiment 1. As described in experiment 1, durational measurements were made using a combination of waveforms, wide-band spectrograms, and pitch tracks generated by Waves+. Criteria for segmentation were the same as those described for Experiment 1.

For each speaker the same 45 items were selected for further measurement. Items were discarded when they were produced with an intonation contour other than the one intended (e.g. a statement instead of a question), or when parts of the utterance were disfluent. For the speakers of Group DG it was not always clear which accent placement the speaker intended to produce. It was therefore decided that for Group DG items were not to be discarded on the basis of nucleus location (but see further section 2.4). However, nucleus location was a criterion for data selection for Group G, that is an item was discarded when it was produced with another nucleus placement than the one intended in the design. The items were selected by the present author and checked by a native speaker of Greek, who agreed in 100% of the cases when deciding which nucleus placement was produced.

As for each speaker the same 45 items were selected, for some speakers there were missing values for some of the items. These missing values were replaced by a subject's mean values for the appropriate stress type, weighted by the mean group values of the item in question. The amount of missing values was less than 5%.

Measurements were made at specific F0 points in the utterance that were thought (after initial visual inspection of the data of both groups) to capture the differences between native and non-native speakers of Greek. Two types of measurement were made, measurements of the scaling (i.e. the fundamental frequency level at which a H or L tone occurs), and measurements of the alignment (in ms) of these F0 points relative to segmental landmarks. The following measurements were made:

- Scaling of H: the highest F0 in the final rise-fall.
- Scaling of L%: the utterance-final low F0.
- Initial pitch direction: the F0 difference of the onset and offset of the stressed vowel of the word bearing the nuclear stress.
- Position of H: the distance (in ms) between the offset of the stressed vowel of the final content word and the H.

#### Statistical design

The statistical design was mixed (within and between items). There were three factors: FINAL WORD STRESS PATTERN, GROUP, and SPEAKER. The factor FINAL WORD STRESS PATTERN is between-items and has three levels (final, penultimate, and antepenultimate). The factor GROUP is within-items and has two levels (DG, and G). The factor SPEAKER is within-items, has eleven levels (DG1, DG2, DG3, DG4, DG5, DG6, G3, G4, G6, G7, G10) and is nested within the factor GROUP. Once general analyses had been completed, separate ANOVAs on the data of each speaker were performed. This was done in order to establish whether all speakers were equally successful in the L2.

### 2.3. Methodology

Before discussing the results, some methodological issues relating to specific problems which were encountered, and which influenced the experimental investigation, need to be discussed. As noted in Chapter 2, the problem of comparing differences in pitch range or scaling across different groups of speakers or languages may be influenced greatly by inter-speaker variation in pitch range. In Figure 4.1, the raw F0 values for each of the speakers of group G and DG are shown on various target points of Greek NNF yes/no questions. The F0 values represent each speaker's F0 values at various target points, averaged over 45 items. From this figure it can be seen that there are considerable inter-speaker differences. That is, although the intonation patterns look roughly the same (with a few exceptions), it is clear that there are differences in range between the different speakers.

As described in the introduction, it was hypothesised that the group of non-native speakers may experience difficulty producing Greek yes/no questions, specifically its final rise-fall. On first inspection of the data for the two groups, it seemed that there was a difference in the height of the final peak between native and non-native speakers. However, because of differences in the pitch range of individual speakers, it was impossible to decide on the basis of raw F0 data, whether these differences are attributable to group differences, or are due to individual speakers' differences in pitch range. For this reason, it was thought necessary to abstract away from these differences between speakers. Therefore, for each of the scaling measurements F0 values were expressed on a speaker-specific normalised scale which was derived by assigning a value of 100 to the top and a value of 0 to the bottom of the speakers' overall F0 range. This approach to normalising F0 data is based on normalising models suggested by Earle (1975) and Rose (1987). A detailed description of the way in which the speaker-specific scales were calculated is given in

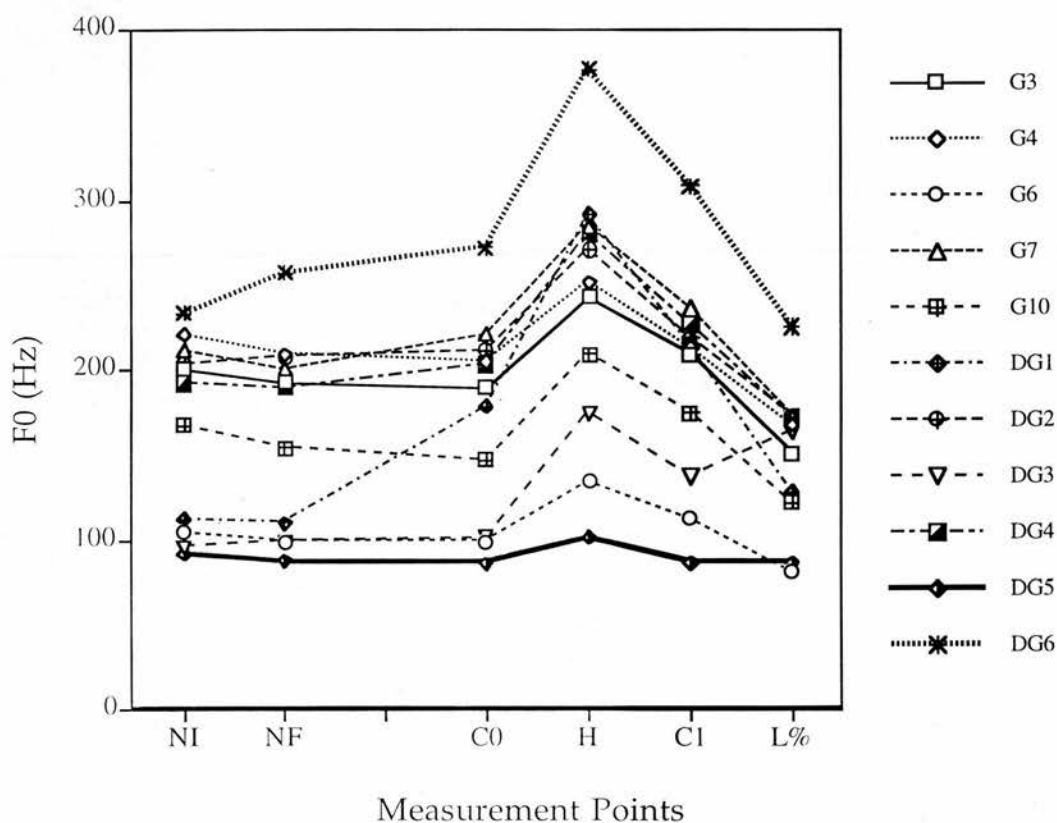


Figure 4.1. Raw F0 values for each of the speakers of Group G and DG on various target points in Greek NNF yes/no questions. The measurement points are: the onset (NI) and offset (NF) of the accented vowel of the first content word; the consonant onset of the stressed syllable of the final content word (C0); the highest F0 in the final rise-fall (H); the offset of the stressed syllable of the final content word (C1); and the utterance-final low F0 (L%).

Appendix D. The effectiveness of this method can be seen in Figure 4.2, which shows the normalised values at the different target points for each of the speakers of Group G. From this figure it becomes clear that most of the transformed F0 values cluster tightly for the speakers of Group G. However, this figure also reveals that speaker G7 differs slightly from the other speakers of her group. In the method section it was mentioned that speaker G7 had a slightly different accent from that of the other speakers, because she was the only speaker who had not been raised in Athens. At the time it was thought that her Pelopponesian accent did not affect her intonational system. However, figure 4.4 reveals that this may not be true, specifically as she differs from the other speakers at the measurement point C0 (the consonant onset of the accented syllable of the final content word), and possibly at L%.

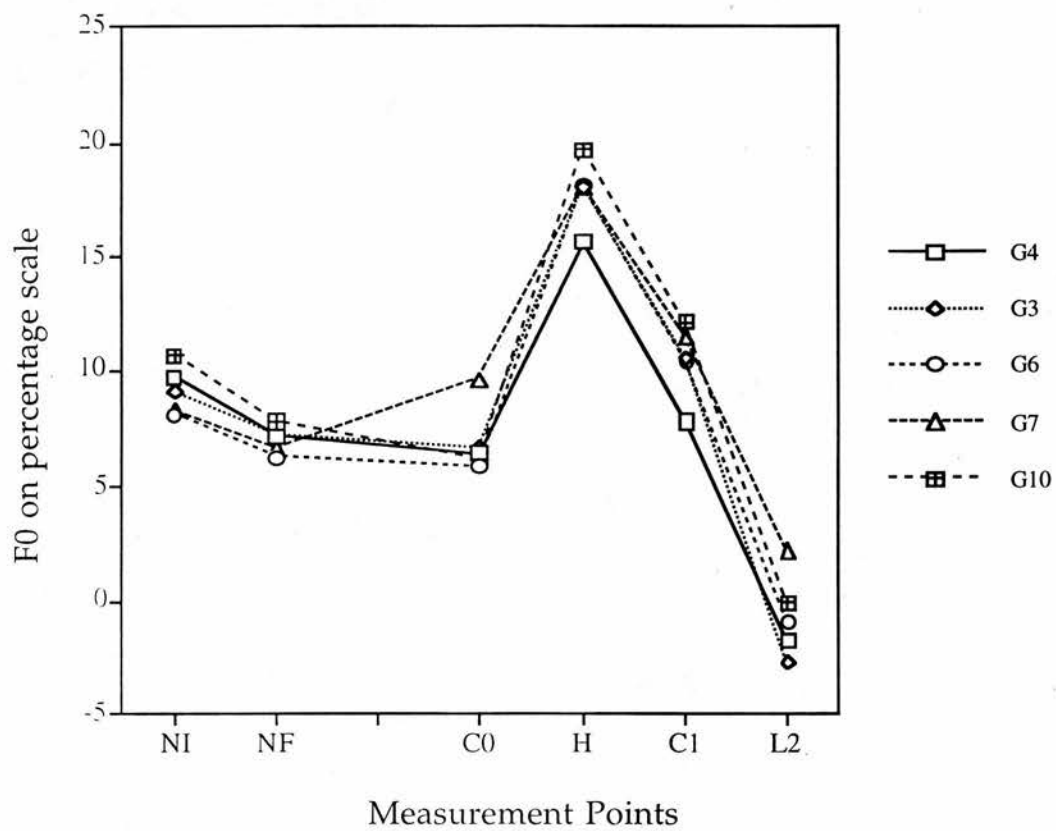


Figure 4.2. Normalised values (expressed at a speaker-specific or percentage scale) at the different target points for each of the speakers of Group G.

## 2.4. Results

### Nucleus placement

As mentioned in the previous section, no explicit instructions were given to speakers about where to place the nuclear accent. Although in nearly all cases (99%) the native Greek speakers put the accent where it was intended in the design of the materials, non-native speakers (as predicted) appeared not to show such high agreement. Therefore, before group analyses were carried out it was established on the basis of the measurement 'position of H' (the interval between the offset of the stressed vowel of the final content word and the H) whether the speakers of Group DG had produced the intended nucleus placement. If speakers had achieved correct nucleus placement, this interval should be negative (since the peak should be somewhere in the accented syllable). This procedure was carried out for each speaker of Group DG, averaged over the three types of final word stress pattern (final, penultimate, and antepenultimate). It should be noted that for the items with word stress on the utterance-final syllable it cannot be established on the basis of the measurement 'position of H' which nucleus placement is produced, as the stressed syllable of the final content word (i.e. the default location of the peak in NNF questions) and the utterance-final syllable (the default location of the peak in the NF questions) in this condition coincide. However, it was thought that by averaging over the three types of word stress, a rough indication could be given about the success achieved, and it was preferred over a more subjective auditory judgement. Table 4.2 shows mean duration (ms) and standard errors of 'position of H' for each of the speakers of Group DG along with the overall means of Group G. From the means it is clear that speakers DG3 and DG5 are not producing the intended nucleus placement, as for them the H occurs long after the stressed syllable.

Table 4.2. Mean duration (ms) of "position of H" (the interval between the offset of the stressed vowel of the final content word and the H) for each of the speakers of Group DG, along with the overall means for Group G. Means are averaged for the three types of final word stress pattern (final, penultimate, and antepenultimate).

NNF	DG1	DG2	DG3	DG4	DG5	DG6	Group G
Mean	-44.33	-90.29	117.93	-75.09	84.51	-64.93	-88.50
SE	8.30	6.98	20.57	6.53	20.59	8.84	10.13

Further inspection of the duration of this interval for each of the items showed that for speaker DG3, as expected, this interval was negative only in the items with word stress on the utterance-final syllable. For speaker DG5 there were two further items in which this interval was negative, but in all other items it was positive. On the basis of these results it seems safe to conclude that neither speaker DG3 nor DG5 had achieved the correct nucleus location for the NNF yes/no questions. Therefore, their results were excluded from further analysis. After this initial data selection, group analyses were carried out for the speakers who had achieved correct nucleus placement, i.e. speakers DG1, DG2, DG4 and DG6.

#### 2.4.1. Group results

##### Scaling of H

In order to establish whether there were differences in the scaling of the H of the utterance-final rise-fall between Group DG and Group G, the means for the dependent variable "scaling of H" (on a percentage scale) were calculated for each item, averaging over the speakers of each group. These means were then entered into a two-way ANOVA (GROUP X FINAL WORD STRESS PATTERN) with repeated measures on the

variable GROUP. The results of this ANOVA show that there is a significant effect of the factor GROUP, indicating that the H is scaled considerably higher for Group DG than for Group G [ $F(1,42)=906.761$ ,  $p<0.0001$ ]. In addition, there was also a significant effect of FINAL WORD STRESS PATTERN [ $F(2,42)=5.466$ ,  $p<0.008$ ], but no significant interaction between factors [ $F(2,42)=3.029$ ,  $p=0.059$ , ns]. Table 4.3 presents the means and standard errors for the scaling of H for each of the groups and stress conditions. The data are also graphed in Figure 4.3.

Table 4.3. Means and standard errors of 'scaling of H' (the peak of the rise-fall expressed on a percentage scale) for the different groups and final word stress patterns.

Final word stress pattern	Group DG		Group G	
	mean	SE	mean	SE
utterance-final	29.68	0.47	16.28	0.66
penultimate	30.39	0.38	18.66	0.62
antepenultimate	29.81	0.40	18.75	0.38

Although the interaction between the factors GROUP and FINAL WORD STRESS PATTERN did not reach significance ( $p=0.59$ , ns), it can be seen from the means and Figure 4.3 that only for Group G the scaling of H is affected by the final word stress pattern. This was confirmed by post hoc tests (Bonferroni). These show that for Group G the H is scaled lower when stress is on the utterance-final syllable, than when it is on the penultimate or antepenultimate syllable of the final content word, between which there is no difference (for final vs penultimate,  $p<0.014$ ; final vs antepenultimate,  $p<0.010$ ; penultimate vs antepenultimate,  $p=1$ , ns). For Group DG there is no significant effect of FINAL WORD STRESS PATTERN (for final vs penultimate,  $p=0.72$ , ns; final vs antepenultimate,  $p=1$ , ns; penultimate vs antepenultimate,  $p=1$ , ns).

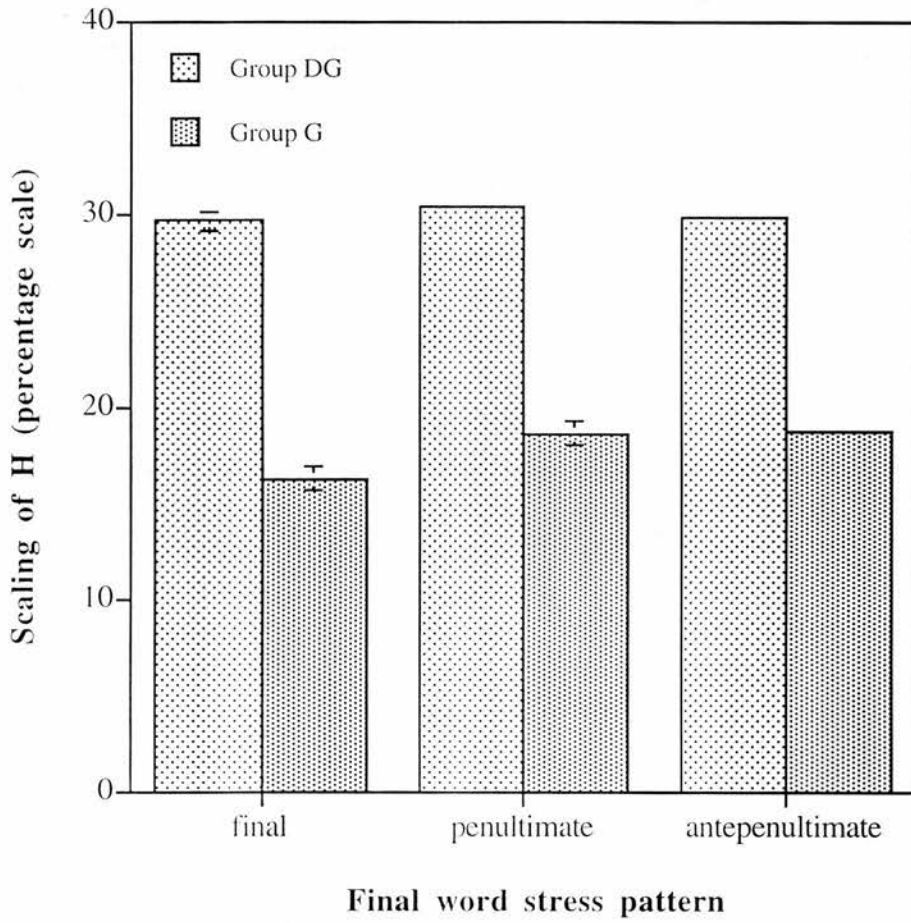


Figure 4.3. Means and standard errors of 'scaling of H' (the peak of the rise-fall expressed on a percentage scale) for the different groups and final word stress patterns. (Missing error bars indicate that the standard errors are very small.)

## Scaling of L%

Table 4.4. Means and standard errors of 'scaling of L%' (the utterance-final low F0 expressed on a percentage scale) for the different groups and final word stress patterns.

Final word stress pattern	Group DG		Group G	
	mean	SE	mean	SE
utterance-final	9.20	0.66	-.69	.33
penultimate	5.61	.35	-.31	.17
antepenultimate	4.54	.34	-1.01	.18

To establish whether there were differences in the scaling of the L% (the utterance-final low F0) between Group DG and Group G, the means for the dependent variable "scaling of L%" (on a percentage scale) were calculated for each item, averaging over the speakers of each group. These means were then entered into a two-way ANOVA (GROUP X FINAL WORD STRESS PATTERN) with repeated measures on the variable GROUP. The results of this ANOVA show that there is a significant effect of the factor GROUP [ $F(1,42)=708.189, p<0.0001$ ], with the L% scaled higher for Group DG. The results also show a significant effect of FINAL WORD STRESS PATTERN [ $F(2,42)=18.366, p<0.0001$ ], and this time there is a significant interaction between the two factors [ $F(2,42)=26.936, p<0.0001$ ]. Post hoc Bonferroni tests reveal that this is due to the fact that the L% is affected by the final word stress pattern in Group DG only. For Group DG the L% is highest when stress is on the utterance-final syllable, lower when stress is on the penultimate or antepenultimate between which there is no difference (for final vs penultimate,  $p<0.0001$ ; for final vs antepenultimate,  $p<0.0001$ , for penultimate vs antepenultimate,  $p=0.352$ , ns). For Group G there is no significant effect of FINAL WORD STRESS CONDITION (for final vs penultimate,  $p=0.819$ , ns; final vs antepenultimate,  $p=1$ , ns; penultimate vs antepenultimate,  $p=1.36$ , ns).

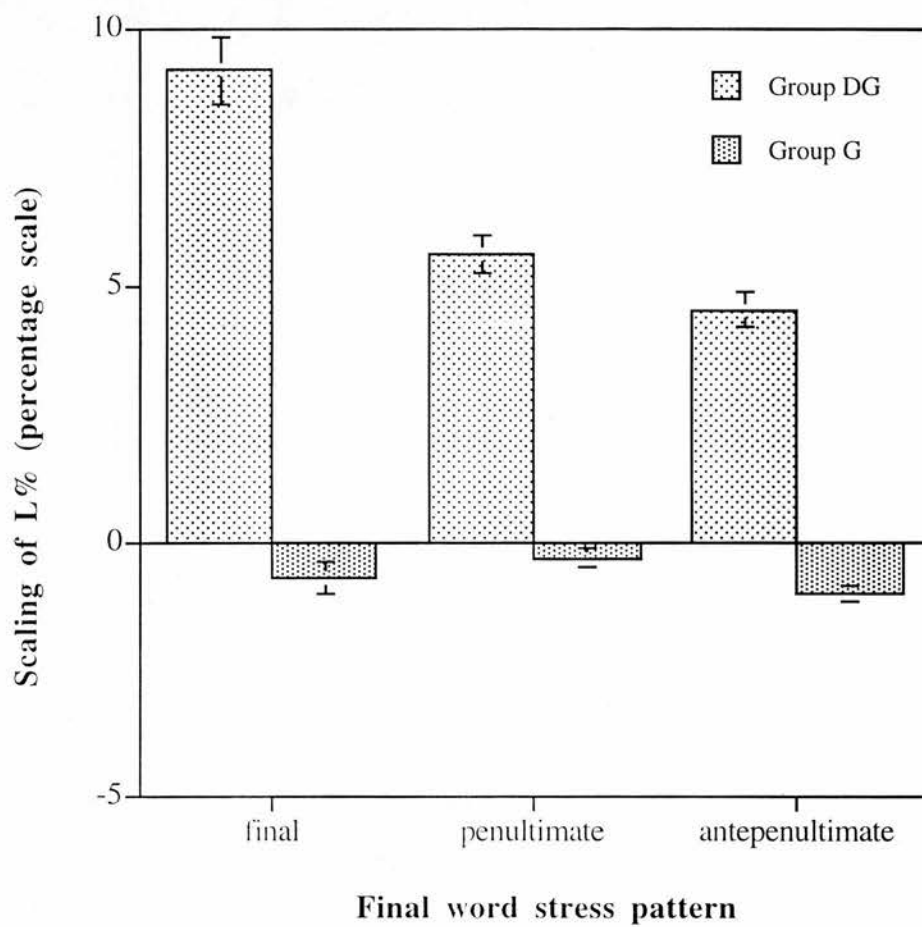


Figure 4.4. Means and standard errors of 'scaling of L%' (the utterance-final low F0 expressed on a percentage scale) for the different groups and final word stress patterns

This is the reverse of what was seen in H scaling, which was affected by the final word stress pattern for Group G but not for Group DG. The means and standard errors for 'scaling of L%' are presented in Table 4.4 and graphed in Figure 4.4.

#### Initial pitch direction

Just as for the other measurements, first the means were calculated for each item, averaging over the speakers of each group. These means of 'initial pitch direction' (the F0 difference between the onset and offset of the stressed vowel of the word bearing the nuclear stress) were then entered into a two-way ANOVA (GROUP X FINAL WORD STRESS PATTERN) with repeated measures on the variable GROUP. The results of this ANOVA show that there is a significant effect of the factor GROUP [ $F(1,42)=101.308, p<0.0001$ ], indicating that pitch is falling for Group G, whereas it is slightly rising for Group DG. Furthermore, the results show that there is no significant effect of the factor FINAL WORD STRESS PATTERN [ $F(2,42)<2.3$ ], and no interaction between factors [ $F(2,42)<1$ ]. The means and standard errors for 'initial pitch direction' (on a percentage scale) are presented in table 4.5.

Table 4.5. Means and standard errors of 'initial pitch direction' (the F0 difference of the onset and offset of the stressed vowel of the word bearing the nuclear stress, expressed on a percentage scale) for the different groups and final word stress patterns.

Final word stress pattern	Group DG		Group G	
	mean	SE	mean	SE
utterance-final	.56	.58	-2.33	.42
penultimate	.55	.39	-2.89	.54
antepenultimate	1.16	.41	-1.46	.16

## 2.4.2. Individual speakers' results

Following the group analyses, separate ANOVAs were performed for the three measurements (scaling of H, scaling of L%, and pitch direction at the nuclear accent) on the data of the individual speakers. The results of these separate ANOVAs, which compared the data of each speaker with the group means of Group G, show that all speakers exhibit the same pattern on all three measurements. Since no immediately interpretable effects of the factor final word stress pattern were found in the group analyses, and it was thought that this factor would not provide further insight about the main purpose of this study (i.e. to determine whether speakers of Group DG had acquired native Greek NNF yes/no question intonation), this factor was left out of the individual analyses.

For 'scaling of H' all speakers show a significant effect of the factor GROUP (for speaker DG1 [ $F(1,42)=853.993, p<0.0001$ ], for speaker DG2 [ $F(1,42)=209.696, p<0.0001$ ], for speaker DG4 [ $F(1,42)=371.918, p<0.0001$ ], and for speaker DG6 [ $F(1,42)=551.449, p<0.0001$ ]). That is, for all speakers the H was scaled significantly higher than that of Group G. The means and standard errors for each speaker together with the group means for Group G are graphed in Figure 4.5.

For 'scaling of L%' the results also show a significant effect of the factor GROUP (for speaker DG1 [ $F(1,42)=108.687, p<0.0001$ ], for speaker DG2 [ $F(1,42)=108.266, p<0.0001$ ], for speaker DG4 [ $F(1,42)=293.264, p<0.0001$ ], and for speaker DG6 [ $F(1,42)=395.528, p<0.0001$ ]). That is, for all speakers the L% is higher than that of Group G. Figure 4.6 shows the means and standard errors for each speaker together with the group means for Group G.

For 'initial pitch direction' the factor GROUP again proved to be significant (for speaker DG1 [ $F(1,42)=21.098, p<0.0001$ ], for speaker DG2 [ $F(1,42)=39.391, p<0.0001$ ], for speaker DG4 [ $F(1,42)=26.748, p<0.0001$ ], and for speaker DG6 [ $F(1,42)=110.985, p<0.0001$ ]). That is, the initial pitch is rising

or only slightly falling for Group DG, whereas it is rising considerably for Group G. The results are graphed in Figure 4.7.

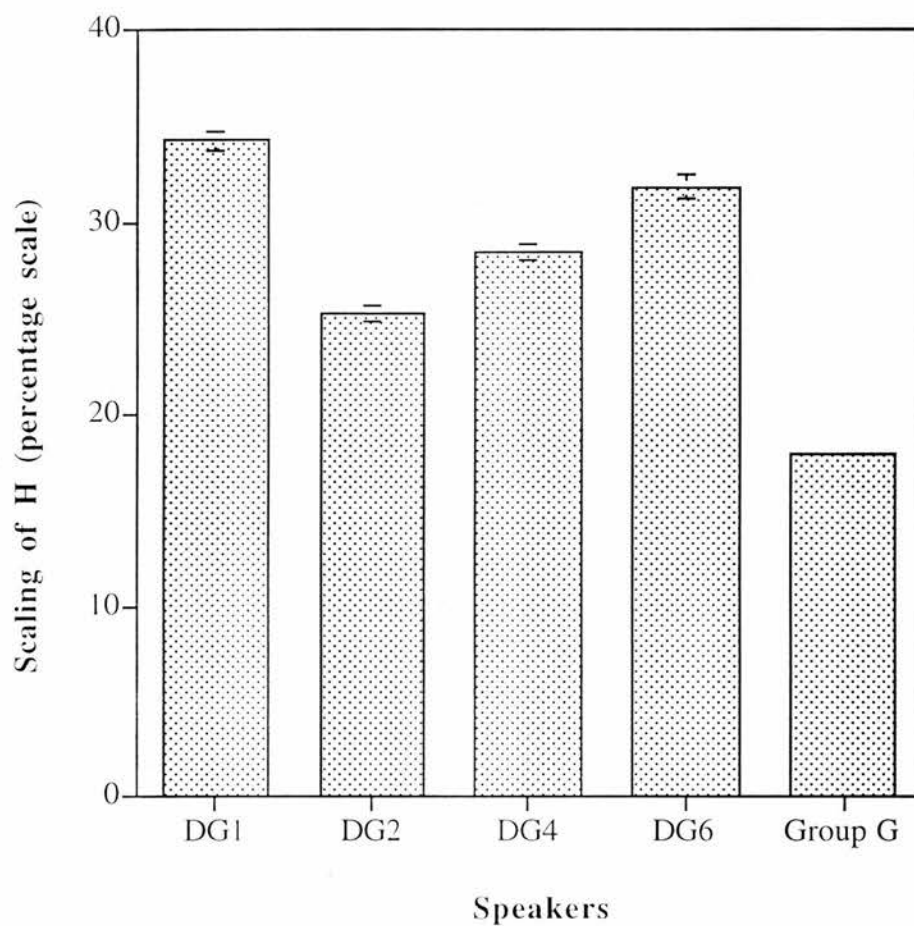


Figure 4.5. Means and standard errors of 'scaling of H' (the peak of the final rise-fall, expressed on a percentage scale) for each speaker of Group DG, together with the group means of Group G.

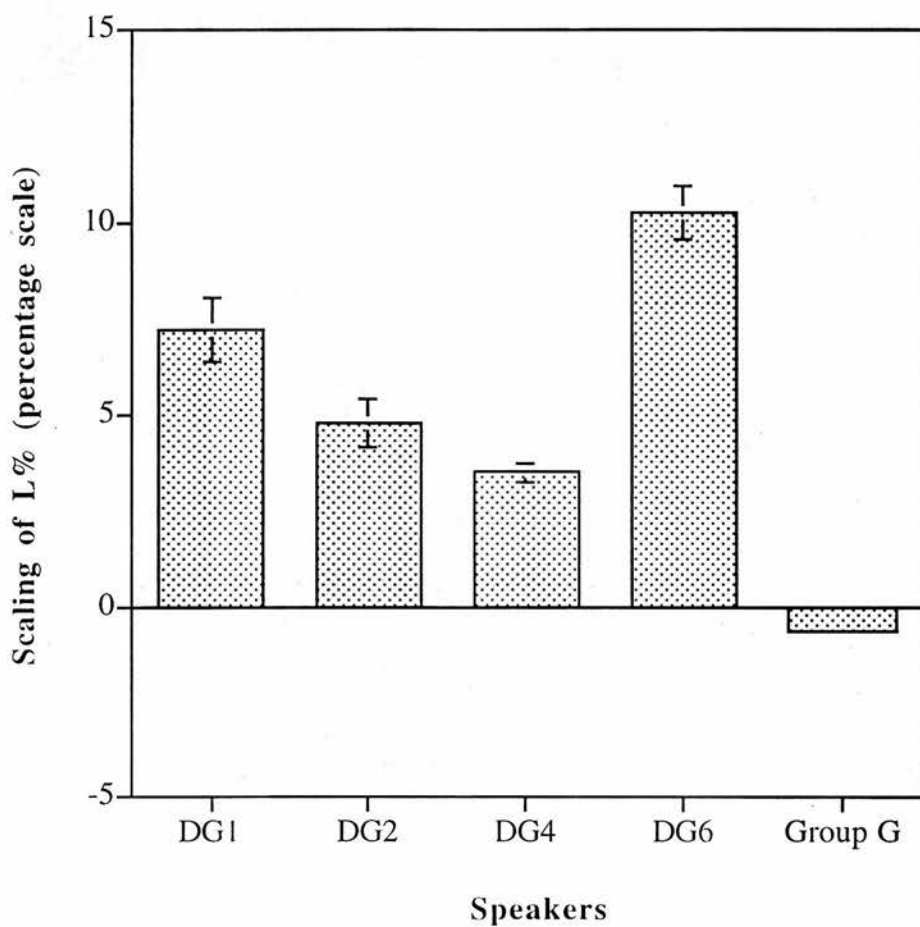


Figure 4.6. Means and standard errors of 'scaling of L%' (the utterance-final low F<sub>0</sub>, expressed on a percentage scale) for each speaker of Group DG, together with the group means of Group G.

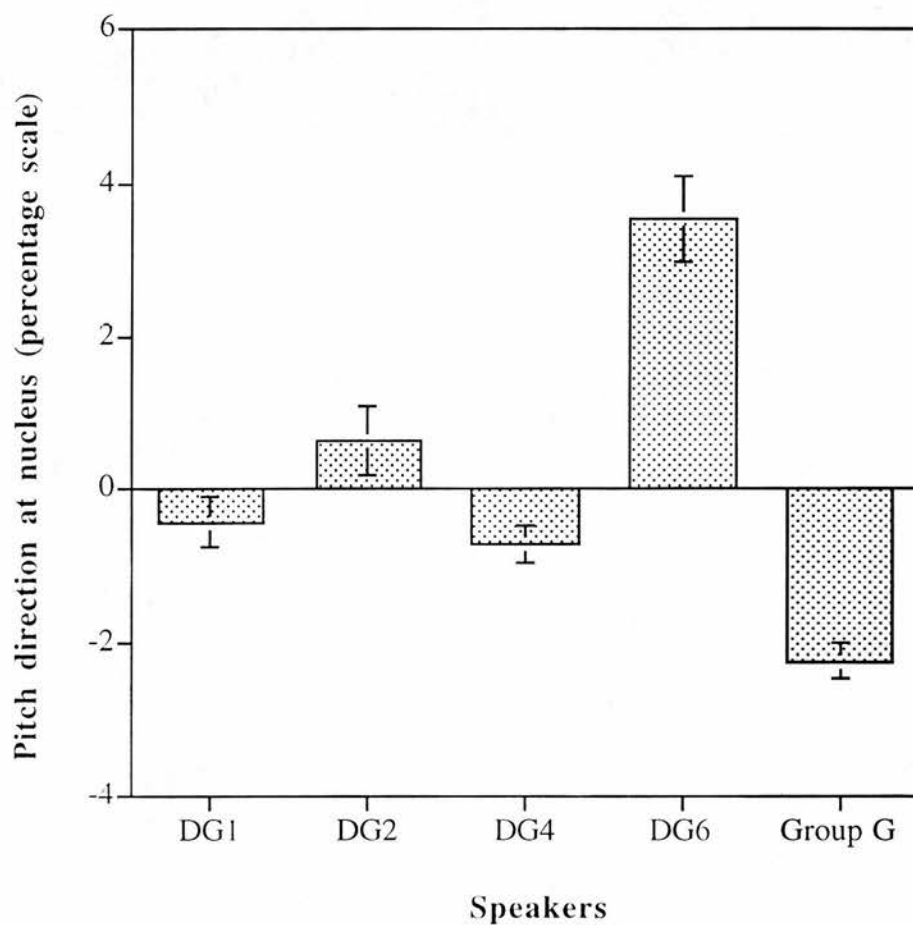


Figure 4.7. Means and standard errors of 'initial pitch direction' (the  $F_0$  difference of the onset and offset of the stressed vowel of the word bearing the nuclear stress, expressed on a percentage scale) for each speaker of Group DG, together with the group means of Group G.

## 2.5. Summary of results

In this experiment, the production of Greek nucleus-non-final (NNF) yes/no questions by non-native (Dutch) speakers of Greek (Group DG) was investigated. Several measurements were taken for both the native (Group G) and the non-native (Group DG) group of speakers. Firstly, it was established on the basis of the durational measurement 'position of H' (the distance (in ms) between the offset of the stressed vowel of the final content word and the peak of the final rise-fall) whether speakers of Group DG had produced the correct nucleus placement. For Greek native speakers it was reported in the literature that in NNF yes/no questions the H is located on the lexically stressed syllable of the final content word (Arvaniti et al., forthcoming). It was therefore hypothesised that the measurement 'position of H' should be negative if the speakers had achieved the correct nucleus placement. It was shown that two of the speakers of Group DG (DG3, DG5) had not achieved correct nucleus placement. Therefore, their results were excluded from further analysis.

Secondly, several scaling measurements (i.e. the fundamental frequency level at which a H or L tone occurs) were taken for the scaling of H (the peak of the final rise-fall) and the scaling of L% (the utterance-final low F0). These raw F0 measurements were converted to a speaker-specific (percentage) scale, which is described in Appendix D.

Then, group analyses were carried out for the speakers who had achieved correct nucleus placement, i.e. means were calculated from the data of speakers DG1, DG2, DG4 and DG6.

The hypothesis tested in this experiment was that Group DG would experience difficulties with the realisation of the 'similar' Greek NNF yes/no questions. These difficulties would result in inaccurate production of this 'similar' contour. In general, this was supported by the experimental results. Specifically, it is shown that for Group DG the H

(the peak of the final rise-fall) is higher than the H produced by native speakers of Greek.

Furthermore, it was shown that the utterance-final low (L%) also appears to be different for the two groups: it is higher for Group DG than for Group G. In fact, Arvaniti et al (forthcoming) had suggested that in Greek yes/no questions the final rise-fall usually falls as low as the final low in statements. It appears that this is not the case for the speakers of Group DG. Their final fall resembles traditional analyses of Greek yes/no question intonation, which describe the contour as a 'fall to mid pitch' (Holton, Mackridge, and Philippaki-Warburton, 1997; Joseph and Philippaki-Warburton, 1987) or 'raised fall' (Mackridge, 1985), and are based on impressionistic data of Waring (1976). It is possible that L2 learners have been influenced by these traditional analyses.

The results also show that the F0 at the stressed syllable of the nuclear accent is rising or slightly falling for the non-native (Dutch) speakers of Greek, whereas it is falling considerably for the native speakers of Greek. It appears that speakers of Group DG have difficulty producing the falling nuclear accent (L\*) of Greek, possibly because pitch accents in Dutch are commonly realised as a 'hat pattern', with rising pitch associated with the lexically stressed syllable (e.g. 't Hart and Cohen, 1973; 't Hart et al., 1990). In native Greek, this L\* accent is rather different, and realised as "a long low (slightly declining) stretch which include[s] the nuclear syllable but show[s] no specific F0 dip associated with it" (Arvaniti et al., forthcoming).

Finally, for the group of non-native speakers there is an effect of final word stress pattern for the L% scaling (with the L% higher when stress is on the utterance-final syllable, than when it is on the penultimate or antepenultimate syllable, between which there is no difference), but not for the H scaling. This is exactly the opposite of what was found for the group of native speakers of Greek. It is not clear how to interpret this difference in the effect of final word stress pattern between the two groups of speakers.

Taken together, the results show that even though four out of the six speakers of Group DG had achieved the correct nucleus placement of the Greek NNF yes/no questions, their production of this contour is inaccurate in all measurements tested: their pitch at the nuclear accent is rising or only slightly falling instead of falling considerably; their scaling of the peak of the rise-fall is too high; and their final rise-fall doesn't fall as low as for speakers of Group G.

### 3. Experiment 5: nucleus final polar questions

In Experiment 4 it was seen how a 'similar' yes/no question contour is realised by non-native (Dutch) speakers of Greek. Experiment 5 will provide data on another Greek yes/no question contour, the NF contour, which as noted in section 1 can be considered 'new'. If applied to intonation, the SLM would predict that when the L2 contour is noticeably different from the L1 contour, the L2 learner will eventually establish a new category for this 'new' L2 contour. Production of this 'new' contour should be unproblematic, although not necessarily produced in the same way as it is by native speakers (for a detailed discussion see Chapter 1, section 4). In any case, it should be less problematic than a 'similar' contour. The goal of the following experiment is to investigate whether the concepts of 'new' and 'similar' can indeed be successfully used to explain L2 intonation data, and to provide instrumental data to support or reject the hypothesis that L2 learners experience more difficulties with a 'similar' than with a 'new' contour. If the latter is true, it is expected that the data from Group DG for the Greek NNF contour will be further removed from the Greek norms than are the data for the NF contour. For this aim, the data from this experiment were compared to the data obtained in Experiment 4.

### 3.1. Method

#### Materials

Sixty Greek yes/no questions were constructed in which the nuclear accent was expected to be placed on the final content word of the sentence, i.e. these were ‘nucleus-final’ (NF) yes/no questions.<sup>4</sup> The sentences were constructed following the same considerations as described for Experiment 4. The materials, as noted in the report of Experiment 4, were constructed with three different types of stress on the final content word (utterance-final, penultimate, and antepenultimate), had at least two unstressed syllables between the stressed syllable of the first and last content word, and used sonorants in the stressed syllable of the final word. As in Experiment 4, the materials were presented in short dialogues in order to achieve the desired placement of the nuclear accent. No instructions were given about where to place the nuclear accent, and again Greek speakers put the accent according to expectation in nearly all cases (99%). Table 4.6 gives some examples of the items used in this experiment. A full list of the test items for nucleus-final (NF) yes/no questions is given in Appendix E.

Table 4.6. *Experiment 5: sample test items for nucleus-final (NF) yes/no questions with different stress patterns of the final content word (adapted from Arvaniti et al (forthcoming). The test word is underlined.*

Nucleus location	Final word stress pattern	Sample test item	Gloss
NF	final	[itane <u>ka'lo</u> ]	“Was it <u>good</u> ?”
	penultimate	[xtipai to <u>ku'duni</u> ]	“Is the <u>bell</u> ringing?”
	ante-penultimate	[xriazetai <u>si'deroma</u> ]	“Does it need <u>ironing</u> ?”

<sup>4</sup> As mentioned before (see previous footnote), the materials were previously used for another experiment and are reported in Arvaniti et al. (forthcoming).

## Procedure

The speakers who participated in Experiment 4 also participated in Experiment 5. The materials for experiment 5, as noted in the report of Experiment 4, were interspersed with the materials for the latter experiment. Digitisation and measurements were the same as for Experiment 4. The only difference was that in Experiment 5, the measurement 'initial pitch direction' did not reflect the F0 difference of the onset and offset of the stressed vowel of the *nuclear word*, but rather of the previous content word (i.e. in both experiments the measurement was taken at the first content word, but in experiment 4 this was the nuclear word, whereas in experiment 5 it was not).

Fifteen items were selected for each speaker following the same selection criteria as in Experiment 4. Missing values were replaced as in Experiment 4. The same normalisation scale was used as the one described in Experiment 4.

## Statistical design

The statistical design was mixed (within and between items). There were three factors: NUCLEUS LOCATION, GROUP, and SPEAKER. The factor NUCLEUS LOCATION is between-items and has two levels (NF and NNF). The factor GROUP is within-items and has two levels (DG. and G). The factor SPEAKER is within-items, has eleven levels (DG1, DG2, DG3, DG4, DG5, DG6, G3, G4, G6, G7, G10) and is nested within the factor GROUP. Once general analyses had been completed, separate analyses were performed on the data of those speakers who only produced the NF yes/no question contour.

### 3.2. Results

#### Nucleus placement

As in Experiment 4, before group analyses were carried out it was established on the basis of the interval 'position of H' (the distance between the offset of the stressed vowel of the final content word and the H) whether the speakers had produced the intended nucleus placement for NF yes/no questions. If the intended nucleus placement was produced, the H should be placed on the final unstressed syllable and 'position of H' should be positive. Table 4.7 shows the mean duration (in ms) of 'position of H' for each of the speakers of Group DG along with the group means for Group G. It is clear that, apart from speaker DG2, all speakers have produced the intended NF contour.

Table 4.7. Mean duration (ms) of "position of H" (the interval between the offset of the stressed vowel of the final content word and the H) in the NF condition for each of the speakers of Group DG along with the overall means for Group G. Means are averaged over the three types of final word stress pattern (final, penultimate, and antepenultimate).

NF	DG1	DG2	DG3	DG4	DG5	DG6	Group G
Mean	103.27	-73.09	121.31	79.73	92.36	95.93	57.54
SE	21.00	11.87	18.67	17.02	20.48	23.39	22.07

When comparing the means of 'position of H' from Experiment 4 with those of Experiment 5 (Table 4.2 vs Table 4.7), it seems clear that some speakers, instead of producing two different focus conditions for Greek yes/no contours (i.e. NNF or NF), just produce one type of focus. Where speaker DG2 only produces the NNF yes/no question contour, speaker DG3 and DG5 produce the NF yes/no question contour only.

Separate ANOVAs for each of these three speakers confirm that for these speakers there is no effect of nucleus location, thus indicating that they only produce one type of focus in yes/no questions (for DG2 [ $F(1,88)=1.559, p=n.s.$ ]; for DG3 [ $F(1,88)=0.015, p=n.s.$ ]; for DG5 [ $F(1,88)=0.073, p=n.s.$ ]).

As the aim of this experiment is to compare NF with NNF conditions for the two groups of speakers, group analyses are based on the three speakers that produced a distinction between the two conditions (i.e. DG1, DG4 and DG6). The other speakers will be analysed separately.

A comparison of Table 4.2 and 4.7 also shows that the standard errors in the former table are considerably larger than those in the latter table. This is just as expected, since in the NF condition the H is expected to occur on the utterance-final syllable, and therefore the distance between the offset of the stressed vowel of the final content word and the H is more variable in NF than in NNF yes/no questions (where the H occurs on the stressed syllable).

### 3.2.1. Group analyses

#### Scaling of H

The data for the H scaling in NF yes/no questions were compared with the scaling data for the H in the NNF yes/no questions from Experiment 4. The aim was to see whether speakers of Group DG experience the same amount of difficulty in the production of scaling of the peak in NNF and in NF yes/no questions. Since Experiment 4 did not show any consistent effects of final word stress pattern, this factor was left out of the analyses in this experiment.

Just as in Experiment 4, first the means were calculated for each item, averaging over the speakers of each group. These means were then entered into a two-way ANOVA (GROUP  $\times$  NUCLEUS LOCATION) with

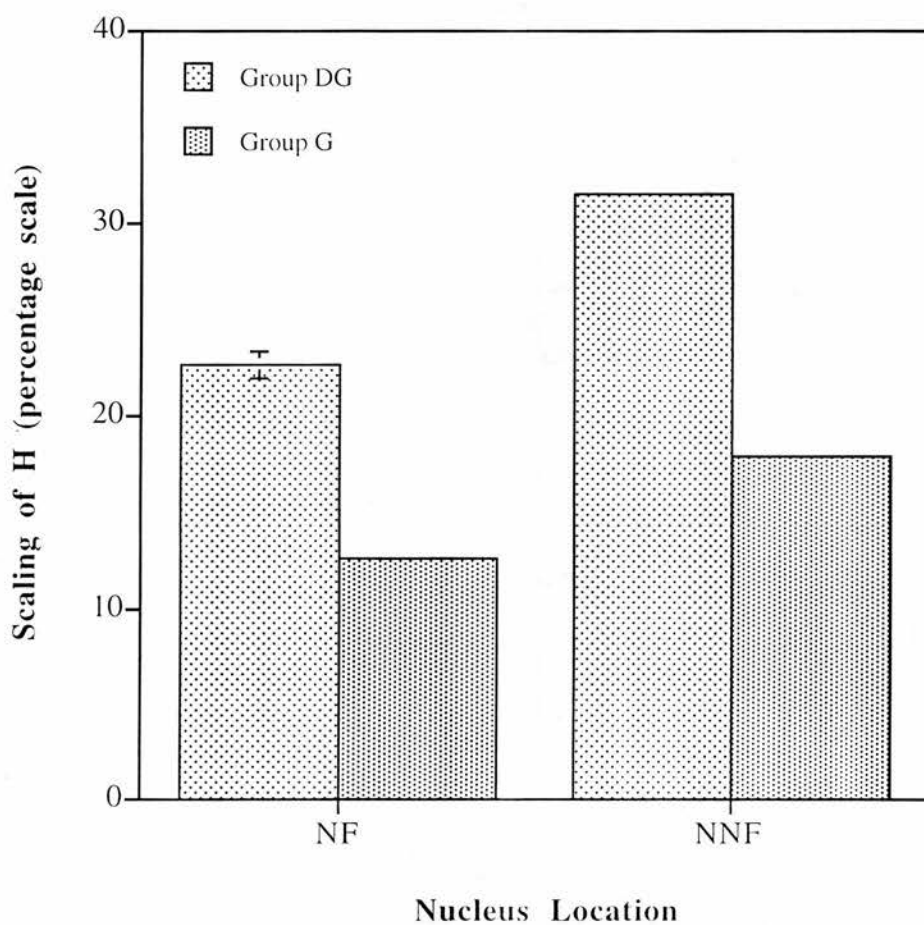


Figure 4.8. Means and standard errors of 'scaling of H' (the peak of the rise-fall expressed on a percentage scale) for the different groups of speakers.

repeated measures on the variable GROUP. Recall that the measurements were expressed on a speaker-specific (percentage) scale.

The results show that there is a significant effect of the factor NUCLEUS LOCATION [ $F(1,88)=190.466, p<0.0001$ ], indicating that the H is scaled higher in the NNF than in the NF condition. In addition, there is also a significant effect of the factor GROUP [ $F(1,88)=1038.972, p<0.0001$ ], with the H scaled higher for Group DG than for Group G. Furthermore, there is an interaction between the two factors [ $F(1,88)=23.040, p<0.0001$ ]. This interaction is just as expected according to the SLM, since Group DG should experience more difficulties (and thus be further removed from the means of Group G) producing the peak in the 'similar' NNF contours than in the 'new' NF contour. Figure 4.8, in which the means and standard errors for 'scaling of H' are graphed, shows that this is exactly the case. The means depicted in Figure 4.8 are pooled for the speakers of each group. As the individual speakers of each group behaved in the same way, no individual means are given for neither of the dependent variables.

#### Scaling of L%

For each item the means of the measurement 'scaling of L%' (on a percentage scale) were calculated, averaging over the speakers of each group. These means were entered into a two-way ANOVA (NUCLEUS LOCATION  $\times$  GROUP), with repeated measures on the factor GROUP. The results show that there is no significant effect of NUCLEUS LOCATION [ $F(1,88)=3.8, p=0.053, ns$ ], but there is an effect of GROUP [ $F(1,88)=653.504, p<0.0001$ ], showing that the L% is scaled higher for Group DG than for Group G. There was no interaction between factors [ $F(1,88)<1, p<0.45, ns$ ]. Figure 4.9 shows the means and standard errors for 'scaling of L%'.

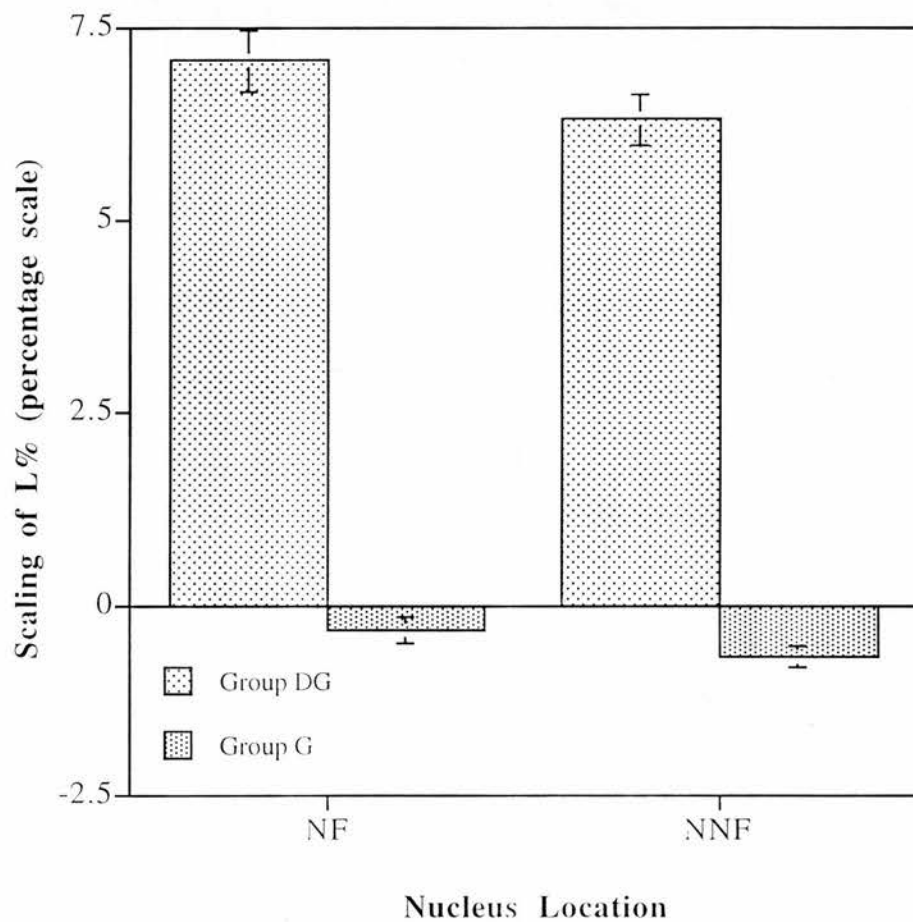


Figure 4.9. Means and standard errors of 'scaling of L%' (the utterance-final low F0, expressed on a percentage scale) for the different groups of speakers.

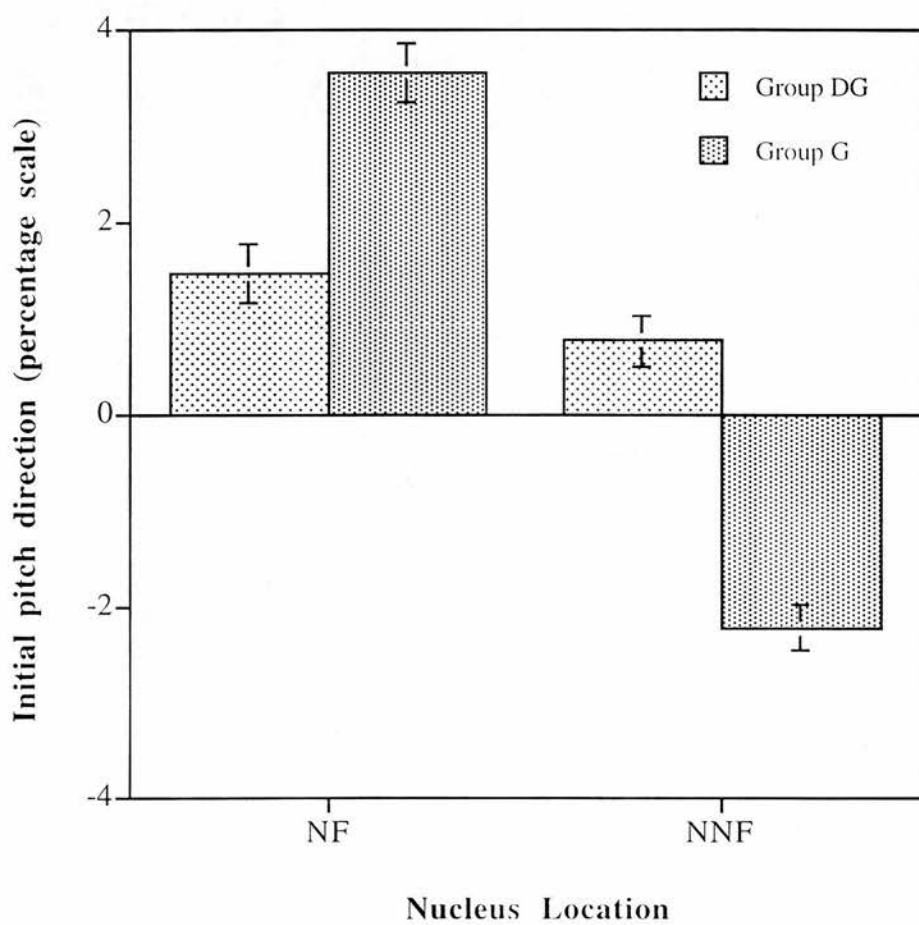


Figure 4.10. Means and standard errors of 'initial pitch direction' (the direction of  $F_0$ , i.e. rising or falling, at the first content word) for the different groups of speakers.

### Initial pitch direction

The initial pitch direction was obtained by subtracting the normalised F0 values at the offset of the stressed vowel of the first content word from those at the onset of the stressed vowel. Recall that for the NNF yes/no questions the first content word was also the word bearing the nuclear accent, but that in NF yes/no questions the nuclear accent was located on the final content word.

The means for each item, averaged over the speakers of each group, were calculated and entered into a two-way ANOVA (NUCLEUS LOCATION  $\times$  GROUP) with repeated measures on the factor GROUP. The results show that the factor GROUP is just significant [ $F(1,88)=4.059$ ,  $p<0.047$ ]. There was also a significant effect of NUCLEUS LOCATION [ $F(1,88)=97.814$ ,  $p<0.0001$ ], and a significant interaction between factors [ $F(1,88)=129.397$ ,  $p<0.0001$ ]. The interaction was due to the fact that for Group DG the pitch is rising in both the NF and the NNF condition, whereas for Group G it is rising in NF condition but falling in NNF condition. Means and standard errors for 'initial pitch direction' are graphed in Figure 4.10.

#### 3.2.2. Individual analyses for speaker DG3 and DG5

In general, the speakers DG3 and DG5 show a similar pattern to the other speakers of Group DG in the NF condition (which is the only nucleus location they produce). For both speakers the H is scaled significantly higher than for Group G (for speaker DG3 [ $F(1,44)=897.217$ ,  $p=0.0001$ ]; for DG5 [ $F(1,44)=72.160$ ,  $p=0.0001$ ]). Furthermore, there is also no significant difference in the pitch direction at the lexically stressed syllable of the first content word of speaker DG3 and Group G, both showing a rising pitch. Speaker DG5, however, shows a different pattern. His pitch direction is significantly different from that of Group G since it

is falling instead of rising [ $F(1,44)=145.717, p=0.0001$ ]. Regarding the L% scaling data, for both speakers the L% is scaled higher than the data of Group G (for DG3 [ $F(1,44)=1902.330, p=0.0001$ ]; for DG5 [ $F(1,44)=57.230, p=0.0001$ ]). Specifically, the L% boundary tone of speaker DG3 is scaled around the same level as his average values for the H (peak of the final rise-fall), suggesting that he has replaced the L% boundary tone by a H% boundary tone.

The results of speaker DG2 are not discussed here, since she only produced the NNF yes/no question contour, and her results have therefore already been discussed in experiment 4.

### 3.3. Summary of results

The results of Experiment 5 show that only half of the non-native (Dutch) speakers of Greek produce the two different nucleus locations (which reflect differences in focus) which are made by native speakers of Greek. Instead, they consistently produce just one of the two possible nucleus locations (NF or NNF). Two speakers (DG3 and DG5) consistently produce NF yes/no questions, whereas one speaker (DG2) consistently chooses the NNF location.

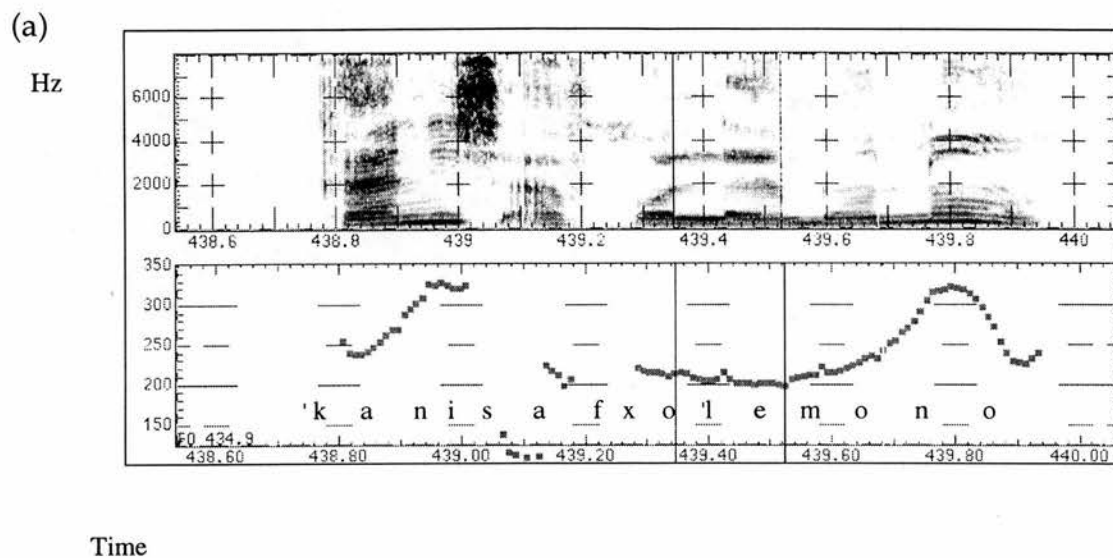
The speakers who did produce a distinction between the two types of nucleus location, were not as consistent in placing the nuclear accent where it was expected as the speakers of Group G. So, speakers of Group DG would sometimes produce a NF yes/no question where speakers of Group G produced a NNF yes/no question (or the other way around). Furthermore, it was shown that the NF contour, which does not have a counterpart in the L1, is more accurately produced than the NNF contour, which is more similar to the L1. This is just as predicted by the SLM. This finding is consistent with findings at the segmental level, like those found by (Flege, 1987a; Flege, 1987b), who showed that American English learners of French pronounced French /y/ (a 'new' sound for English

speakers) more accurately than /u/ (a 'similar' sound for English speakers).

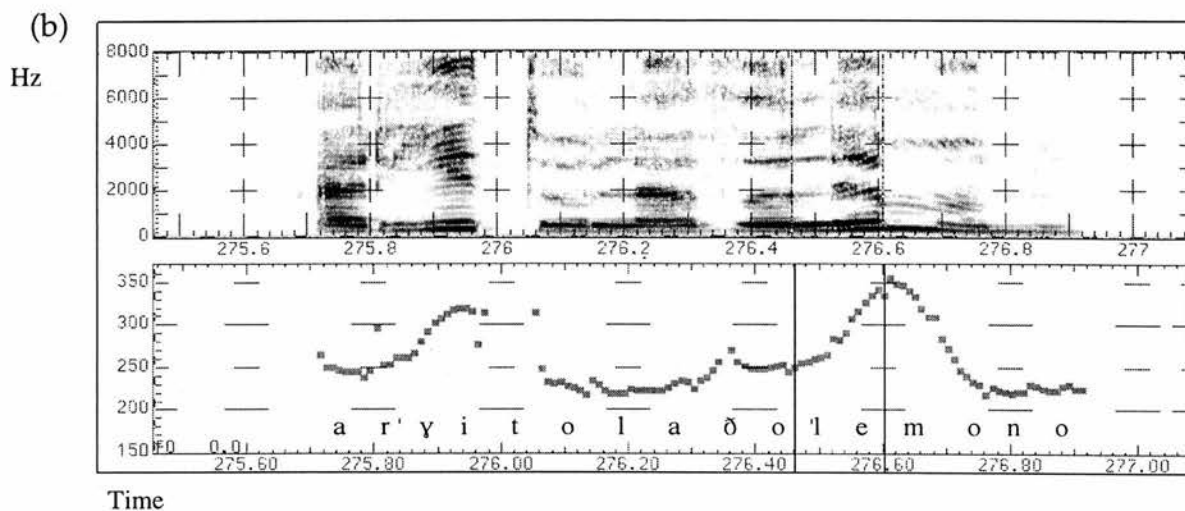
However, the results also show that it was not only the 'similar' yes/no question contour which was inaccurately produced by the non-native speakers, but also the 'new' yes/no question contour. This was not exactly as predicted by the SLM, which posits that when an L2 sound is 'new' it should be possible to establish a new category for this sound. Therefore, production and perception of this sound should be fairly unproblematic.

The differences between the native (Group G) and non-native speakers (Group DG) are manifested in all the measurements tested. Specifically, it was shown that the peak of the final rise-fall as well as the L% boundary tone are higher for Group DG in both the NF and the NNF yes/no questions. Furthermore, the F0 direction at the stressed vowel of the syllable which bears nuclear stress (in the NNF condition) is rising for Group DG, whereas for Group G it is falling. Finally, the data for Group DG show that the obtained values for the scaling of the H (the peak of the final rise-fall) are closer to the native Greek norm in the new NF yes/no questions than in the similar NNF yes/no questions.

Figure 4.11 illustrates how Greek yes/no questions are produced by a speaker of Group DG. For clarity, the same questions (but now produced by a speaker of Group G) are shown in Figure 4.12 (repeated from figure 2.11).



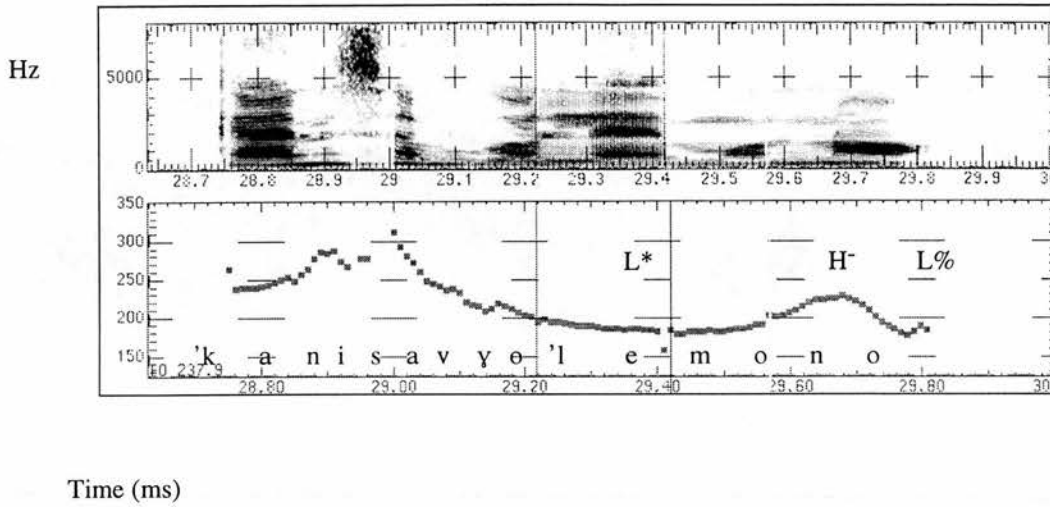
Time



Time

Figure 4.11. An example of a NF (top panel) and NNF (bottom panel) yes/no question read by one of the speakers of Group DG (DG6). The vertical lines delimit the beginning and end of the stressed syllable of the final content word.

(a)



(b)

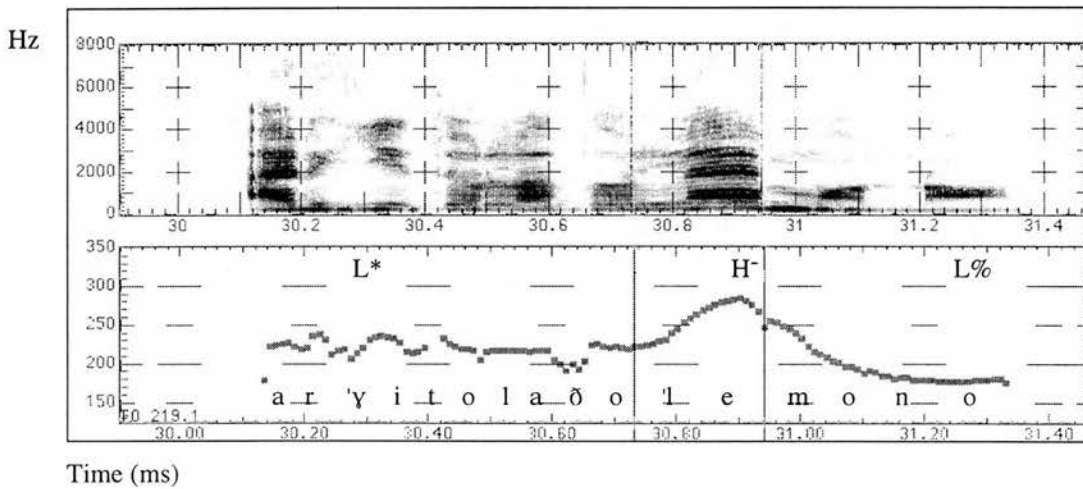


Figure 4.12. The same sentences as those shown in Figure 4.13, but now read by a speaker of Group G (G4). The vertical lines delimit the beginning and end of the stressed syllable of the final content word.

#### 4. Discussion

This chapter has presented experimental evidence bearing on the intonation of yes/no questions produced by non-native (Dutch) speakers of Greek. In particular, it was found that none of the L2 speakers has reached native values for any of the measurements tested. This finding is contradictory to the findings of a previous experiment (Experiment 2, Chapter 3), where it was found that one speaker had achieved native values for the timing of the peak in Greek prenuclear rising accents. This suggests that not all aspects of L2 intonation constitute the same amount of difficulty for L2 learners. Perhaps it is easier to master temporal aspects of L2 speech. This is consistent with findings on the perception of L2 segments (as described in Jenkins and Yeni-Komshian, 1995), who suggest that temporal aspects of L2 speech phenomena, like VOT, are easier to acquire than are other phenomena, like spectral change or place contrasts.

A striking difference which was found between Group DG and Group G was the use of the two different nucleus locations. In the majority of cases the speakers of Group G put the nuclear accent where it was intended in the design of the materials. The speakers of Group DG, on the other hand, showed far less agreement. Only half of the speakers produced the two different nucleus locations. The other half consistently produced just the NNF location (speaker DG2), or only the NF location (speakers DG3 and DG5). It is not clear whether this is due to the fact that they cannot produce such a distinction, or rather that they fail to recognise from the supplied dialogues which nucleus placement is required. Perhaps they have failed to understand the difference in focus between Greek and Dutch. As mentioned before, the neutral nucleus location in Greek yes/no questions is on the verb, whereas in Dutch having the main accent on the verb would make the question non-neutral. It is possible that speakers of Group DG have failed to recognise this difference in focus, and interpret the two different nucleus

placements as two different nuclear accent types. That is, they interpret the NF and the NNF yes/no question as two different question tunes, both of which have the nuclear accent on the final content word. This would be hardly surprising, since Dutch also has (at least) two different yes/no question tunes. This would explain why speakers at this proficiency level (with on average 17 years of experience with the L2) still fail to produce a distinction which was produced in most cases by the native speakers of Greek.

Besides the fact that they experienced difficulties with the actual choice of nucleus location, all speakers of Group DG also failed to produce the desired contours in phonetically the same way as did Greek native speakers. Both NNF and NF yes/no question contours were produced differently by the two groups at several identifiable points in the utterances. First of all, there was a difference in the scaling of the H (the peak of the final rise-fall), with the H higher for Group DG than for Group G. Secondly, there was a difference in the scaling of the L%. The L% was scaled higher for Group DG than for Group G. Thirdly, in NNF yes/no questions the pitch direction at the first content word (which in this case was the word bearing the nuclear accent) was rising for Group DG, whereas it was falling for Group G. In NF yes/no questions, however, the pitch direction at the first content word (which now did not bear the nuclear accent) was rising for both groups. However, there was a small difference between the extent of the rise, with the rise slightly larger for Group G than for Group DG.

Taken together, the results of this series of experiments seem to support the position put forward in Flege's SLM (e.g. Flege, 1995), that L2 learners will experience more difficulties producing a 'similar' rather than a 'new' L2 target<sup>5</sup>. Although originally developed to account for segmental data, the results of experiment 4 and 5 suggest that the predictions the SLM makes about the degree of difficulty of L2 sounds,

---

<sup>5</sup> The term 'target' will be used here to refer to both segmental and suprasegmental aspects of L2 speech.

may be applied to aspects of L2 prosody. However, it should be noted that the SLM's predictions are only upheld on the assumption that the acoustic-phonetic similarity was judged correctly. In the introduction it was mentioned that the notion of acoustic-phonetic similarity is rather ill-defined, and that it was not straightforward to make the distinction between 'new' and 'similar' yes/no questions. It was decided to determine the similarity of the Greek yes/no question contours to Dutch contours on the basis of the location of the final rise-fall. It was argued that the NNF yes/no question contour, which has a rise-fall which is associated with a stressed syllable, is similar to Dutch. The NF contour, with a rise-fall associated with a boundary, was thought to be new for Dutch speakers. However, there are many more differences between Dutch and Greek yes/no question intonation (as mentioned in Chapter 2, section 3), which may influence the perceptual distance between the L1 and L2. If these differences were taken into account when determining the degree of similarity between the two languages, predictions may have been different. In Chapter 5, the problem of determining phonetic similarity and the consequences this has for the predictions the SLM makes will be discussed further.

There is one remaining finding which cannot be fully explained by the SLM. Although the current data indicate that L2 learner's production of the new NF contour is closer to the native Greek norm than that of the similar NNF contour, speakers of Group DG have still failed to produce the new contour accurately. This difficulty with the production of a new contour is not the expected result, as according to the SLM, it should be possible to establish a new category for a new L2 target and production of this new target should therefore be unproblematic. This result implies that the SLM's classification of phonetic similarity may be rather problematic (at least for intonational data), and that equivalence qualification is not readily predictable. This problem will also be further discussed in Chapter 5.

## Chapter 5. General discussion

### 1. Overview of the main findings

The instrumental work reported in this thesis can be divided into two main areas: one which is concerned with pitch alignment in Greek prenuclear accents (chapter 3), and another with the scaling and nucleus placement in Greek yes/no questions (chapter 4). In both areas of research the production of these sentence types by non-native (Dutch) speakers of Greek was examined.

#### 1.1. Alignment

The alignment studies consisted of three experiments. The first experiment was a pre-test to investigate the alignment of the L and H tone in Dutch prenuclear accents, and consequently to establish whether there are cross-linguistic differences in alignment between Greek and Dutch prenuclear accents. It was found that Greek and Dutch do not differ in their alignment of the L in prenuclear accents. In both languages the L is located at the onset of the accented syllable of the word bearing the prenuclear accent. The experiment revealed, however, that the prenuclear accents peak earlier in Dutch than in Greek, confirming that there are indeed cross-linguistic differences in peak alignment between the two languages. The results also revealed that there is a difference in the alignment of the peak in Dutch depending on whether the vowel is long or short. If the vowel is phonologically long, the peak is located near the end of the accented vowel, but if it is short the peak occurs on the post-vocalic consonant.

This finding (that there are cross-linguistic differences in peak alignment between Dutch and Greek) paved the way for the following two experiments, which investigated whether non-native (Dutch) very

advanced speakers of Greek (Group DG) are able to produce native-like peak alignment values in the L2 (experiment 2), and whether the acquisition of an L2 has any effect on their peak alignment production in the L1 (experiment 3).

The data from experiment 2 suggest that four of the five speakers of Group DG have not acquired native peak alignment values in the L2. However, it was also found that they - contrary to prediction based on Flege's Speech Learning Model (SLM) - have not developed a merged system intermediate between the L1 and L2. Instead, it seems that they transfer the peak alignment of their L1 to the L2. Specifically, it was concluded that they place the peak as early as they would do in the L1 when the accented vowel is phonologically long.

Although most of the speakers failed to acquire native values of Greek peak alignment, one speaker (DG4) managed to produce peak alignment values which were within the norm for native speakers. As speaker DG4 was by no means the most experienced L2 speaker this was not the expected result. On the basis of the data available no firm conclusions could be made about the reason for her success. However, it was pointed out that she differed from the other speakers in age of first exposure to the L2 and in the learning situation, and it was suggested that this could possibly have contributed to her success. In section 2.3 the implications of this success on assumptions about limits of language learning will be further discussed.

The final alignment experiment (experiment 3) found that experienced L2 learners show a different pattern of peak alignment in their L1 (Dutch) than do native L1 speakers who do not have extensive experience in an L2 (Group D). In experiment 1, it was found that all speakers of Group D showed an effect of the phonological length of the accented vowel of the test word on the alignment of the peak. For Group DG this effect was much smaller, and was not observed in the data of all speakers. As there were no obvious differences between the two groups apart from experience with the L2, it was suggested that this may have

caused the speakers of Group DG to show a different alignment pattern in the L1.

Experiment 3 also showed that the speaker who had achieved native L2 peak alignment values (Speaker DG4) also shows native values of peak alignment in the L1. That is, her Dutch prenuclear accents peak earlier when the accented vowel of the prenuclear test word is long than when it is short. In other words, her peak alignment production shows native values in both the L1 and the L2. This is contrary to the expectation put forward in the SLM, which posits that the better L2 learners become at approaching the L2 norm the more their L1 will be affected.

The final conclusions made were concerned with the nature of the observed bidirectional influence (L2 influencing L1, and vice versa). No evidence of such bidirectional influence was found for speaker DG4, as she maintained native values of peak alignment in both the L1 and the L2. For the other speakers it was concluded that their alignment patterns were collapsed. Most of them produced just one alignment category for all three conditions (i.e. Greek, Dutch short, and Dutch long). That is, the peak in their L2 was as early as that in the Dutch long condition, and they mostly did not produce a distinction between Dutch long and Dutch short peak alignment.

## 1.2. Nucleus placement and scaling

The study on scaling and nucleus placement consisted of two experiments, both of which investigated the production of Greek yes/no questions with different nucleus locations by speakers of Group DG. The first experiment in this series (experiment 4) was concerned with Greek yes/no questions in which the nucleus was located on the first content word of the utterance, i.e. nucleus-non-final (NNF) yes/no questions. The second (experiment 5) compared production of NNF yes/no questions

with that of yes/no questions in which the nucleus occurred on the utterance-final word, i.e. nucleus-final (NF) yes/no questions.

In experiment 4, it was found that the speakers of Group DG, even though they are very experienced in the L2, show inaccurate production of NNF yes/no questions. This is the expected result as it was predicted (based on the SLM) that even advanced speakers of the L2 would fail to establish a new category for a 'similar' NNF yes/no question contour, and therefore their production of this contour was expected to be inaccurate. Inaccurate production was found for all the investigated scaling measurements, which are in general scaled higher for Group DG than for the group of Greek native speakers (Group G).

In experiment 5, it was found that only half of the speakers of Group DG produce the two different nucleus locations which are made by speakers of Group G. Furthermore, those speakers that do produce a distinction produce the NF contour more accurately than the NNF contour. This is congruent with the SLM's assumption that it is easier to learn a novel L2 target (like the NF yes/no question contour) than it is to learn to produce a similar target (like the NNF yes/no question contour). However, it was found that Group DG has not only failed to produce the similar NNF yes/no question contour accurately, but also the novel NF contour. Specifically, it was found that the peak and final low are scaled higher for Group DG than for group G (but less so in NNF than in NF yes/no questions). This finding goes against the SLM's position that novel L2 targets should not constitute any problem for L2 learners. This problem will be further discussed in the discussion section.

Finally, it was found that speaker DG4 was not as successful in producing L2 scaling in yes/no questions as in the peak alignment in prenuclear accents. Although she managed to achieve native L2 peak alignment, she failed to exhibit native values for any of the scaling measurements obtained from the yes/no questions. Perhaps not all aspects of L2 intonation constitute the same amount of difficulty for L2 learners.

## 2. General discussion

One of the aims of this thesis was to investigate whether the SLM - one of two influential models of L2 phonetic learning which have been developed to account for segmental aspects of L2 acquisition - can also account for prosodic aspects of L2 acquisition. In general it was found that many of the predictions the SLM would make (if applied to prosodic data) were upheld. However, there were several problematic results which indicated that the model needs to be further refined, and specifically that the unit of analysis may not be appropriate. It is suggested here that this refinement is not only needed to accommodate prosodic aspects of L2 learning, but may also be appropriate to better account for segmental learning.

### 2.1. Phonetic similarity and the unit of analysis

One of the shortcomings of the SLM and most studies in L2 phonology is that they typically focus on segment sized phenomena. At the time when the contrastive analysis hypothesis was prevalent, L2 difficulties were predicted on the basis of differences between the L1 and L2 at the level of the phoneme. Although Lado (1957) was aware that the position of the phoneme in the syllable and its relation to other adjacent sounds influences production and perception of these phonemes in the L2, this view was not widely adopted in L2 studies until much later. Even in the current models of L2 phonology like Flege's SLM and Best's PAM, the focus remains on segmental sounds. However, the position of a segment within the syllable and its phonetic context is now explicitly recognised by the SLM (and less explicitly by the PAM). The SLM now states that the appropriate level of analysis is the 'position-sensitive allophonic level, rather than [the] more abstract phonemic level' (Flege, 1995: p. 239). Even so, the largest unit examined in virtually all research in the area of L2 phonology is the word, and most research does not go beyond the syllable

sized unit. Thus it seems that most researchers in the field have opted for the allophonic level as the most appropriate level of analysis (cf. Jenkins and Yeni-Komshian, 1995), and as a result other candidates do not get the attention they possibly deserve. There are, however, studies that suggest that units beyond the syllable or word level may be more appropriate, and there are many errors in L2 phonology which cannot be accounted for without taking the possibility of transfer from prosodic rules and/or phonological organisation into account.

Some of the results of the experiments described in this thesis, for example, suggest that it is not sufficient to only look at phonetic similarity in order to adequately predict the degree of difficulty of certain L2 targets. In experiments 4 and 5 described in this thesis, predictions were made about the degree of difficulty Dutch L2 learners would experience in producing the different focus readings (NNF or NF) of Greek yes/no questions. These predictions were based on the SLM, which classes L2 sounds as new or similar on the basis of the phonetic difference between the L1 and L2 sounds. The degree of similarity between Dutch and Greek yes/no questions was therefore described in acoustic-phonetic terms. It was argued that if one compares the phonetic shape of Dutch and Greek yes/no questions, it can be seen that they are fairly similar, in the sense that both languages can have rising-falling accents. However, the position of the rise-fall is crucial in determining whether we are dealing with a similar or novel intonation pattern. In Greek yes/no questions the association of the  $H^-$  phrase accent varies with the location of the nucleus. If the nuclear accent is on the last word (NF), the  $H^-$  occurs on the *utterance-final syllable*, but if it is on an earlier word (NNF) it occurs on the *stressed syllable of the final word*. Since in Dutch a rise-fall is always prominence lending and cannot occur on an unstressed syllable, it transpires that the NNF yes/no question should be seen as similar, and the NF as new.

However, if we restricted ourselves to phonetic similarity and compare the phonetic shape of yes/no questions, we would fail to

recognise some differences which in my opinion are crucial in explaining the difficulty L2 learners experience with Greek yes/no questions. First of all, it is important to realise that the peak of the final rise-fall in Greek yes/no questions is not the nuclear accent, and that native Greek speakers feel that the main accent is on the low-pitched lexically stressed syllable of the preceding word (e.g. Ladd, 1996; Arvaniti et al., forthcoming). In Dutch, however, a rise-fall is prominence lending, and it seems likely that Dutch listeners will hear the final rise-fall as signaling a nuclear accent. So, even though phonetically Greek and Dutch yes/no questions are similar, phonologically they are rather different. Where the peak in Dutch yes/no questions is nuclear, in Greek it is not since it occurs *after* the nuclear accent. It seems therefore likely that Dutch non-native speakers of Greek fail to identify the Greek peak as a post-nuclear accent. Instead, it is likely that the Greek yes/no question is interpreted as a L prenuclear accent followed by a H\* nuclear accent.

In the light of this, it now becomes clear why the group of non-native speakers failed to produce *both* the NNF and the NF yes/no question contour accurately. In both cases, it was found that the H was scaled higher than for the group of Greek native speakers, even though for NF yes/no questions this was not predicted by the SLM. It seems then that in order to determine similarity *both* the phonetic shape *and* the phonological organisation of L1 and L2 intonation patterns need to be taken into account.

A further complicating factor in the production of these yes/no questions, may have been the fact that the default nuclear accent in Greek is located on the verb, whereas in Dutch having the main accent on the verb would make the question non-neutral. It is very well possible that Dutch speakers of Greek do not recognise this difference in focus, and interpret the two nucleus locations as two different question tunes.

The SLM does not make any predictions about this difficulty Dutch speakers of Greek may experience, as it is not concerned with aspects of function and meaning. Nevertheless, as the results of experiment 5 show,

only half of the non-native speakers actually produced a distinction between NNF and NF yes/no questions, which suggests that this distinction was somehow problematic. From my own experience I know that it is difficult for advanced learners to recognise that there is a difference in the relative prominence of Greek NF and NNF yes/no questions. From my previous work it can be seen that even though I am a near-native speaker of Greek and I had had approximately 12 years of experience with the L2 at the time, I wrongly classified Greek NF yes/no questions in which the main stress was placed on the noun as neutral questions (Mennen, 1993).

As such difficulties would not be predicted on the basis of phonetic differences between the two languages under consideration alone, it may be beneficial for research in L2 phonology to not restrict itself to segmental similarity. There are many L2 errors which cannot be accounted for without making reference to units above the syllable or word level. To give just one example, if one listens to Italian speakers of

---

English, one may observe that examples like (a) and (b) below<sup>1</sup>, are typically pronounced in the same way (i.e. with the underlined n at the beginning of the second word). English speakers, on the other hand, would pronounce the n at the beginning of the second word in (a), but at the end of the first word in (b).

(a) Norma Nelson

(b) Normann Elson

Such difficulties are due to the transfer of Italian syllabification rules which allow resyllabification across words. That is, Italian resyllabifies a consonant from the end of one word to the beginning of a following vowel initial word. The fact that rules which apply across words can be transferred, suggests that the word level is too small a unit to account for all L2 errors. Evidence for transfer of syllabification rules have been found amongst others in the English of native speakers of Italian (Vogel, 1991), and in the Arabic of native speakers of English (Broselow, 1988). Flege's SLM obviously falls short in accounting for this kind of data, but it seems clear from the above examples that there may be a role for phonological/prosodic organisation in determining the similarity between L1 and L2 phonology.

## 2.2. Bidirectional influence

Another issue which deserves attention here is the claim that the L1 and L2 interact with one another. That is, not only can the L1 influence the L2 (the traditional interpretation of interference), the L2 may also have an effect on the L1. Researchers have sought to explain this phenomenon by positing that both L1 and L2 phonetic categories are stored together in a 'common phonological space' (Flege, 1995: p. 239, 242).

---

<sup>1</sup> The examples are taken from Ladd and Schepman (work in progress).

This provides an elegant explanation for the finding that experienced L2 learners would merge the properties of similar L1 and L2 sounds.

However, the evidence in support of this assumption may not be that compelling. Most evidence has come from studies which investigated voice onset time (VOT) data (e.g. Caramazza et al., 1973; Williams et al., 1979; Flege and Hillenbrand, 1984). The results of these studies clearly indicate that there are considerable differences between VOT production in L2 learners and native speakers. However, their evidence is based on VOT production data from L2 learners, which is compared to mean VOT values for native speakers obtained from other sources. No attempt has been made to balance the groups which are compared for factors like education, social factors etc. Furthermore, the reference values mentioned in these sources, are sometimes based on rather limited samples, and may not be an accurate reflection of language-specific VOT values (Docherty, 1992). It would have been better if data from comparable groups were used in these studies (for a similar view see Markham, 1997).

The data from the experiments on Greek peak alignment (Chapter 3) are rather confusing with regards to the existence of a common phonological space. On the one hand, most non-native speakers did not even approach native values of peak alignment in the L2. Nevertheless, their L1 peak alignment was not native-like, which supports the assumption of a common (or at least partly overlapping) phonological system. On the other hand, one of the speakers showed evidence of separate categories of peak alignment in the L1 and L2, which suggests that she managed to keep the L1 and L2 phonological systems separate. Perhaps it is this separation of phonological systems which is the key to successful (i.e. native) performance in both the L1 and the L2. If one is able to separate the two languages at the phonological level - as is possibly the case in bilingual children acquiring both languages as infants or young children (e.g. Holm, Ozanne, and Dodd, 1997; Johnson and Lancaster, 1998) - successful performance in both the L1 and L2 should be possible. If, however, a common phonological system is used for both the

L1 and L2 - as is thought to be the case in most L2 learners who started learning in adulthood - this should result in inaccurate production of both the L1 and L2. In fact, this is exactly as most studies suggest: the earlier learning starts the better the chances that learning is successful. Although this is an appealing explanation, it fails to explain why some L2 learners who started learning later in life seem to have separate phonological systems for both languages. In order to answer this more research is needed to test whether there is indeed a difference between the phonological systems of child and adult bilinguals, and whether the age of learning has an effect on the separateness of the phonological system.

### 2.3. Limits on attainment

Some researchers take the view that there are limits to the level of competence an L2 learner can achieve, and posit that it is impossible to produce native-like pronunciation in the L2, at least not at the cost of losing nativeness in the L1 (e.g. Scovel, 1969; Flege, 1995; Major, 1990). However, in the light of the findings described in Chapter 3, this assumption may have to be revised, as it was found that it is possible to reach nativeness in some aspects of L2 prosody (without this having an effect on the L1). In fact, native-like performance has also been found for L2 segments, but curiously this fact most often only gets mentioned in passing and is often not backed up by empirical evidence. Furthermore, as many studies on L2 acquisition report grouped results only, the fact that some individuals do achieve native-like values may be obscured. For example, Markham (1997) observed that although Flege (1991) presented group results for the production of L2 English stops by Spanish learners, it could be determined that some of the Spanish speakers (two or three out of ten) had actually produced VOT values which fell within the native English range.

This implies that there may be many more 'high-achievers' than usually assumed. On the basis of the evidence to date it is impossible to determine why some learners achieve the L2 norm and others fail to do so. On the other hand, there is an abundance of evidence for the existence of factors (such as age, amount and type of training, amount of L1 and L2 input, as described in Chapter 1) which may influence the ability to perform well in another language. Perhaps one should assume that every L2 learner has the potential to achieve nativeness in the L2, but that this potential is inhibited by a combination of these factors (cf. Markham, 1997). In any case, it is an interesting fact that there are exceptions on both ends of the scale, exceptionally good learners on the one hand and exceptionally bad learners on the other. Perhaps we need to concentrate more on these exceptional learners in order to find out what factors play a crucial role in the ability to perform well in a second language.

#### 2.4. Types of phonetic categories

One of the factors that may affect speech production and perception in the L2 is the type of phonetic categories. In Chapter 1 it was suggested that not all phonetic categories constitute the same amount of difficulty in L2 learning. It was suggested that temporal aspects of L2 speech phenomena, such as VOT, may be easier to acquire than are spectral change or place contrasts. The empirical studies described in this thesis also suggest that L2 learners may not perform in the same way on different aspects of L2 intonation. The speaker who achieved native values for L2 peak alignment - which in fact is a temporal aspect - did not manage to achieve native values for the scaling of pitch in yes/no questions. It may be that it is easier to adjust temporal aspects of L2 intonation than it is to adjust the relative level at which a H or a L tone occurs. However, since in this study only one speaker achieved native values for one of the aspects tested, more studies are required to come to any firm conclusions.

### 3. Suggestions for future research

Some of the questions posed in this thesis were answered by the findings of the experimental studies. However, other questions remain unanswered and new questions have arisen out of the results. In this section, some suggestions for future research will be given that can possibly address some of the unanswered questions.

First of all, most readers of this thesis will by now have asked themselves whether the differences which were found between the native and non-native speakers are in fact perceived by Greek native speakers. As the experiments in this thesis only concentrated on the production of L2 intonation, this question cannot be answered without further research. Future perception studies in which native speaker's utterances are manipulated separately for the factors peak alignment and scaling would reveal whether such manipulations are perceived by native speakers. However, these perception experiments are not as straightforward as it may seem. The fact that listeners may perceive early alignment as strange or unacceptable does not necessarily imply that early alignment contributes to the impression of non-nativeness. In fact, it may be more likely that native listeners do not perceive early alignment as such on individual accents. However, this difference in alignment may be perceived globally and it may contribute to the impression of non-nativeness.

Another question, which was touched upon in this thesis but could not be fully answered, is the issue of whether there is a common phonological space for L1 and L2 phonetic categories. Although it is an attractive view which intuitively makes sense, at the moment there is not enough evidence to support it. Some research seems to suggest that there may be a difference between child bilinguals (bilinguals acquiring both languages as infants or young children) and adult bilinguals (who started learning the L2 in adulthood), that is child bilinguals may have two separate phonological systems, whereas adult bilinguals may have a

common phonological system for both L1 and L2. In order to address the issue of separateness of the L1 and L2 phonological systems it may be useful to compare production of certain (supra)segments by child and adult bilinguals in both languages, with that of monolingual native speakers of each language. To give an example, one could investigate the production of peak alignment by Greek/English bilinguals, or the production of VOT by Thai/English bilinguals. In each case, the two groups of speakers should be tested in both their L1 and L2, and for each language the results need to be compared with those of monolingual speakers. If the aforementioned studies are correct, the obtained measurements for the adult bilinguals in both the L1 and the L2 should be different when compared to monolingual values. By contrast, the measurements from the child bilinguals in both languages should be similar to those of the monolingual speakers.

The issue of separateness of phonological systems may not only be important for a better understanding of L2 learning, but it may also have clinical implications. If therapy is given to child bilinguals with phonological disorders, it should only be effective in the language in which it is given (since the child supposedly has two separate phonological systems). Such an effect was indeed found by Holm et al. (1997) who presented a case study of a bilingual child who improved considerably in one of the languages in which therapy was given, but showed no improvement in the other.

The empirical work reported in this thesis focussed on production data obtained from carefully designed materials which were read out by the native and non-native speakers. It would be interesting to acquire more knowledge about the production abilities of non-native speakers in more natural conversational data. It seems likely that speakers' production varies depending on the task they are confronted with. Conversational data may therefore give us a different picture about the speaker's ability to produce L2 data.

Eventually, it is important to extend any theoretical model of speech learning to include a role for prosody. In such a model the notion of phonetic similarity needs to be further refined, and the appropriate unit of analysis needs to be determined. However, before such a model can be developed further empirical data need to be gathered investigating prosodic aspects of L2 learning.

## References

Abramson, A.S., and L. Lisker. (1970). Discriminability along the voicing continuum: Cross-language tests. *Proceeding of the 6th International Congress of Phonetic Sciences*:569-573.

Adams, C., and R. Munro. (1978). In search of the acoustic correlates of stress: Fundamental frequency, amplitude, and duration in the connected utterances of some native and nonnative speakers of English. *Phonetica* 35:125-156.

Anderson-Hsieh, J., R. Johnson, and K. Koehler. (1992). The relationship between native speaker judgments of nonnative pronunciation and deviance in segmentals, prosody, and syllable structure. *Language Learning* 42:529-555.

Archibald, J. (1992). Adult abilities in L2 speech: Evidence from stress. In *New sounds 92: Proceedings of the 1992 Amsterdam Symposium on the acquisition of Second-Language Speech*, edited by J. Leather and A. James. Amsterdam: University of Amsterdam.

Arvaniti, A. (1998). Phrase accents revisited: comparative evidence from standard and Cypriot Greek. *International Conference on Spoken Language Processing (Sydney)*.

Arvaniti, A., and D.R. Ladd. (1995). Tonal alignment and the representation of accentual targets. *Proceedings of the thirteenth international congress of phonetic sciences* 4:220-223.

Arvaniti, A., and D.R. Ladd. (forthcoming). Acoustic properties of tonal targets in Modern Greek WH-question intonation. *Journal of the Acoustical Society of America*.

Arvaniti, A., D.R. Ladd, and I. Mennen. (1998). Stability of tonal alignment: the case of Greek prenuclear accents. *Journal of Phonetics* 26 (1).

Arvaniti, A., D.R. Ladd, and I. Mennen. (in press). What is a starred tone? Evidence from Greek. In *Papers in Laboratory Phonology*, edited by M. Broe and J. Pierrehumbert. Cambridge: Cambridge University Press.

Arvaniti, A., D.R. Ladd, and I. Mennen. (forthcoming). Quantitative F0 descriptions require reference to local targets: the intonation of Greek polar questions and emphatic statements. *Journal of the Acoustical Society of America*.

Backman, N. (1979). Intonation errors in second language pronunciation of eight Spanish speaking adults learning English. *Interlanguage Studies Bulletin* 4 (2):239-266.

Backman, N.E. (1977). Intonation problems of eight Spanish-speaking adults learning English. PhD, Applied Psycholinguistics, Boston University School of Education, Boston.

Beckman, M.E. (1986). *Stress and Non-stress Accent.*, Netherlands Phonetic Archives. Dordrecht: Foris.

Beebe, L.M. (1987). Myths about interlanguage phonology. In *Interlanguage phonology: The acquisition of a second language sound system*, edited by G. Ioup and S. H. Weinberger. Cambridge, Mass.: Newbury House.

Best, C., T., and W. Strange. (1992). Effects of phonological and phonetic factors on cross-language perception of approximants. *Journal of Phonetics* 20:305-330.

Best, C.T. (1995). A direct realist view of cross-language speech perception. In *Speech perception and linguistic experience: Issues in cross-language research*, edited by W. Strange. Timonium MD: York Press.

Birdsong, D. (1992). Ultimate attainment in second language acquisition. *Language* 68:706-755.

Bohn, O.-S. (1995). What determines the perceptual difficulty encountered in the acquisition of nonnative contrasts? *Proceedings of the fourteenth international congress of phonetic sciences* 4:84-91.

Bohn, O.-S., and J.E. Flege. (1990). Interlingual identification and the role of foreign language experience in L2 vowel perception. *Applied Psycholinguistics* 11:303-328.

Bolinger, D. (1978). Intonation across languages. In *Universals of human language*, edited by J. Greenberg. Palo Alto, CA: Stanford University Press.

Bongaerts, T. (1997). Exceptional learners and ultimate attainment in second language acquisition. In *Studies in English language teaching*, edited by J. Aarts, I. De Mönnik and H. Wekker. Amsterdam/Atlanta: Rodopi.

Bongaerts, T., B. Planken, and E. Schils. (1995). Can late learners attain a native accent in a foreign language? A test of the critical period hypothesis. In *The age factor in second language acquisition*, edited by D. Singleton and Z. Lengyel. Clevedon, England: Multilingual Matters.

Bongaerts, T., C. Van Summeren, B. Planken, and E. Schils. (1997). Age and ultimate attainment in the pronunciation of a foreign language. *Studies in second language acquisition* 19:447-465.

Booij, G. (1995). *The phonology of Dutch*. Oxford: Clarendon Press.

Botinis, A. (1989a). Discourse intonation in Greek. *Working Papers* 35, Department of Linguistics, University of Lund:5-23.

Botinis, A. (1989b). Stress and prosodic structure in Greek: A phonological, acoustic, physiological and perceptual study. PhD, Phonetics Department, Lund University, Lund.

Boves, L., B.L. Ten Have, and W.H. Vieregge. (1984). Automatic transcription of intonation in Dutch. In *Intonation, accent and rhythm*, edited by Gibbon and Richter. Berlin/New York: De Gruyter.

Brazil, D. (1985). *The communicative value of intonation in English*. Vol. 8, *Discourse analysis monograph*: University of Birmingham, Department of English.

Broselow, E. (1988). Prosodic phonology and the acquisition of a second language. In *Linguistic Theory in Second Language Acquisition*, edited by S. Flynn and W. O'Neill. Dordrecht: Kluwer.

Brown, G., K. Currie, and J. Kenworthy. (1980). *Questions of intonation*. London: Croom Helm.

Brown, G., and G. Yule. (1983). *Discourse analysis*. Cambridge: Cambridge University Press.

Bruce, G. (1977). *Swedish word accents in sentence perspective*. Lund: Gleerup.

Bruce, G. (1982). Developing the Swedish intonation model. In *Working Papers*. Lund: Department of Linguistics and Phonetics, University of Lund.

Bruce, G., and E. Gårding. (1978). A prosodic typology for Swedish dialects. In *Nordic prosody*, edited by E. Gårding, G. Bruce and R. Bannert: Gleerup.

Buysschaert, J. (1990). Learning intonation. In *New Sounds 90: Proceedings of the 1990 Amsterdam Symposium on the Acquisition of Second-language Speech*, edited by J. Leather and A. James, Amsterdam: University of Amsterdam.

Caramazza, A., G. Yeni-Komshian, E. Zurif, and E. Carbone. (1973). The acquisition of a new phonological contrast: The case of stop consonants in French-English bilinguals. *Journal of the Acoustical Society of America* 54:421-428.

Caspers, J. (1994). Pitch movements under time pressure: Effects of speech rate on the melodic marking of accents and boundaries in Dutch. PhD, Rijksuniversiteit Leiden, Leiden.

Caspers, J. (1998a). Experiments on the meaning of two pitch accent types: the 'pointed hat' versus the accent-lending fall in Dutch. *International Conference on Spoken Language Processing (Sydney)*.

Caspers, J. (1998b). Who's next? The melodic marking of question versus continuation in Dutch. *Language and Speech*.

Caspers, J., and V.J. Van Heuven. (1993). Effects of time pressure on the phonetic realization of the Dutch accent-lending pitch rise and fall. *Phonetica* 50 (3):161-171.

Catford, J.C., and D.B. Pisoni. (1970). Auditory vs. articulatory training in exotic sounds. *Modern Language Journal* 54:477-481.

Cohen, A., and J. 't Hart. (1967). On the anatomy of intonation. *Lingua* 19:177-192.

Collier, R. (1970). The optimum position of prominence lending rises. *IPO Annual Progress Report* 5:82-85.

Collier, R. (1972). From pitch to intonation. PhD, Katholieke Universiteit Leuven.

Collier, R. (1975). Perceptual and linguistic tolerance in intonation. *IRAL* XIII (4):293-308.

Collier, R., and J. 't Hart. (1981). *Cursus Nederlandse Intonatie*. Amsersfoort: Acco.

Coppieters, R. (1987). Competence differences between native and fluent nonnative speakers. *Language* 63:544-73.

Corder, S.P. (1978). Language-learner language. In *Understanding second and foreign language learning: Issues and approaches*, edited by J. Richards. Rowley, Mass.: Newbury House.

Cruttenden, A. (1992). Review of 't Hart, Collier, and Cohen 1990. *Journal of Linguistics* 28 (2):522-526.

Cruz-Ferreira, M. (1983). Non-native comprehension of intonation patterns in Portuguese and in English. PhD, University of Manchester, Manchester.

Cruz-Ferreira, M. (1987). Non-native interpretive strategies for intonational meaning. In *Sound Patterns in Second Language Acquisition*, edited by A. James and J. Leather. Dordrecht: Foris Publications.

Crystal, D. (1995). *A dictionary of linguistics and phonetics*. 3 ed. Oxford: Blackwell Publishers Ltd.

Daan, J. (1938). Dialect and pitch-pattern of the sentence. *Proceedings of the 3rd International Conference of Phonetic Sciences*:473-481.

Dauer, R.M. (1980). Stress and rhythm in Modern Greek. PhD, Department of Linguistics, University of Edinburgh, Edinburgh.

De Bot, K. (1986). The transfer of intonation and the missing data base. In *Crosslinguistic influences in second language acquisition*, edited by E. Kellerman and M. Sharwood Smith. New York: Pergamon Press.

Docherty, G.J. (1992). *The timing of voicing in British English*. Berlin: Foris Publications.

Dulay, H., and M. Burt. (1974). Natural sequences in child second language acquisition. *Language Learning* 24:37-53.

Earle, M.A. (1975). An acoustic phonetic study of Northern Vietnamese tones, Speech Communication Research Laboratory, Santa Barbara.

Eckman, F.R. (1987). On the naturalness of interlanguage phonological rules. In *Interlanguage phonology: The acquisition of a second language sound system*, edited by G. Ioup and S. H. Weinberger. Cambridge, Mass.: Newbury House.

Ellis, R. (1994). *The study of second language acquisition*. Oxford: Oxford University Press.

Es, G.A. Van. (1932). Syntaxis en Dialectstudy IV: Intonatie en Syntaxis. *Onze Taaltuin*:122-128.

Flege, J.E. (1980). Phonetic approximation in second language acquisition. *Language Learning* 30:117-134.

Flege, J.E. (1981). The phonological basis of foreign accent: a hypothesis. *TESOL Quarterly* 15 (4):443-455.

Flege, J.E. (1984). The detection of French accent by American listeners. *Journal of the Acoustical Society of America* 76 (3):692-707.

Flege, J.E. (1987a). Effects of equivalence classification on the production of foreign language speech sounds. In *Sound patterns in second language acquisition*, edited by A. James and J. Leather. Dordrecht: Foris Publications.

Flege, J.E. (1987b). Limits on phonetic accuracy in foreign language speech production. In *Interlanguage phonology: The acquisition of a second language sound system*, edited by G. Ioup and S. H. Weinberger. Cambridge, Mass.: Newbury House.

Flege, J.E. (1987c). The production of "new" and "similar" phones in a foreign language: evidence for the effect of equivalence classification. *Journal of the Acoustical Society of America* 15:47-65.

Flege, J.E. (1991). Age of learning affects the authenticity of voice-onset-time (VOT) in stop consonants produced in a second language. *Journal of the Acoustical Society of America* 89 (1):395-411.

Flege, J.E. (1992). Speech learning in a second language. In *Phonological Development: Models, Research, and Application*, edited by C. Ferguson, L. Menn and C. Stoel-Gammon. Timonium, MD: York Press.

Flege, J.E. (1995). Second language speech learning: Theory, findings and problems. In *Speech perception and linguistic experience: Issues in cross-language speech research*, edited by W. Strange. Timonium MD: York Press.

Flege, J.E., and O.-S. Bohn. (1989). An instrumental study of vowel reduction and stress placement in Spanish-accented English. *Studies in Second Language Acquisition* 11:35-62.

Flege, J.E., E.M. Frieda, and T. Nozawa. (1997). Amount of native-language (L1) use affects the pronunciation of an L2. *Journal of Phonetics* 25:169-186.

Flege, J.E., and J. Hillenbrand. (1984). Limits on phonetic accuracy in foreign language speech production. *Journal of the Acoustical Society of America* 76 (3):708-721.

Flege, J.E., M. Munro, and I. Mackay. (1995). Factors affecting strength of perceived foreign accent in a second language. *Journal of the Acoustical Society of America* 97:3125-3134.

Fokes, J., and Z.S. Bond. (1989). The vowels of stressed and unstressed syllables in nonnative English. *Language Learning* 39 (3):341-373.

Fries, C. (1945). *Teaching and learning English as a foreign language*. Ann Arbor, MI: University of Michigan Press.

Goldsmith, J. (1976). Autosegmental phonology. PhD, MIT.

Goto, H. (1971). Auditory perception by normal Japanese adults of the sounds "l" and "r". *Neuropsychologia* 9:317-323.

Govaert, G.A., and A. Van Katwijk. (1968). Prominence as a function of the location of pitch movement. *IPO Annual Progress Report* 2:115-117.

Greasly, V.M. (1971). An experiment to investigate two methods of teaching French pronunciation to native American speakers of English. PhD, Columbia University, New York.

Grønnum, N. (1991). Prosodic parameters in a variety of regional Danish standard languages. *Phonetica* 47:188-214.

Grosjean, F. (1989). Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain and Language* 36:3-15.

Grover, C., D.G. Jamieson, and M.B. Dobrovolsky. (1987). Intonation in English, French and German: perception and production. *Language and Speech* 30 (Part 3):277-296.

Gussenhoven, C. (1984). *On the grammar and semantics of sentence accents*. Dordrecht: Foris Publications.

Gussenhoven, C. (1988). Adequacy in intonation analysis: The case of Dutch. In *Autosegmental studies on pitch accent*, edited by H. Van der Hulst and N. Smith. Dordrecht-Cinnaminson: Foris Publications.

Gussenhoven, C. (in press). The boundary tones are coming: on the nonperipheral realisation of boundary tones. In *Papers in Laboratory Phonology*, edited by M. Broe and J. Pierrehumbert. Cambridge: Cambridge University Press.

Gussenhoven, C., and J.G. Blom. (1978). The perception of prominence by Dutch listeners. *Phonetica* 35:216-230.

Gussenhoven, C., and T. Rietveld. (1997). Empirical evidence for the contrast between L\* and H\* in Dutch rising contours. *Proceedings of the ESCA Workshop on Intonation: Theory, Models and Applications*:18-20.

Gussenhoven, C., and T. Rietveld. (forthcoming). The behaviour of H\* and L\* under variations in pitch range in Dutch rising contours. .

Haan, J. (1999). Mannen, vrouwen en vragen: verschillen in toonbereik. In *3<sup>e</sup> sociolinguïstische conferentie*, edited by E. Huls and B. Weltens.

Haan, J., V.J. van Heuven, J.J.A. Pacilly, and R. van Bezooijen. (1997). An anatomy of Dutch question intonation. In *Linguistics in the Netherlands*, edited by H. De Hoop and J. Coerts. Amsterdam: Benjamins.

Halliday, M.A.K. (1967). Notes on transitivity and theme in English, Part 2. *Journal of Linguistics* 3:199-244.

Halliday, M.A.K. (1970). *A course in spoken English: intonation*. Oxford: Oxford University Press.

Hardcastle, W.J. (1982). Constraints on coarticulatory processes. In *Linguistic controversies: Essays in linguistic theory and practice in honour of F.R. Palmer*, edited by D. Crystal. London: Edward Arnold.

Hecht, B.F., and R. Mulford. (1987). The acquisition of a second language phonology: interaction of transfer and developmental factors. In *Interlanguage phonology: The acquisition of a second language sound system*, edited by G. Ioup and S. H. Weinberger. Cambridge, Mass.: Newbury House.

Hermes, D., and H.H. Rump. (1994). Perception of prominence in speech intonation induced by rising and falling pitch movements. *Journal of the Acoustical Society of America* 96:83-92.

Hermes, D., and J. Van Gestel. (1991). The frequency scale of speech intonation. *Journal of the Acoustical Society of America* 90:97-102.

Holm, A., A. Ozanne, and B. Dodd. (1997). Efficacy of intervention for a bilingual child making articulation and phonological errors. *International Journal of Bilingualism* 1 (1).

Holton, D., P. Mackridge, and I. Philippaki-Warbuton. (1997). *Greek: a comprehensive grammar of the modern language*. London: Routledge.

Ioup, G. (1995). Evaluating the need for input enhancement in post-critical period language acquisition. In *The age factor in second language acquisition*, edited by D. Singleton and Z. Lengyel. Cleveland, England: Multilingual Matters.

Ioup, G., E. Boustagui, M. El Tigi, and M. Mosell. (1994). Reexamining the critical period hypothesis: A case study in a naturalistic environment. *Studies in second language acquisition* 16:73-98.

Jenkins, J.J., and G.H. Yeni-Komshian. (1995). Cross-language speech perception: Perception and promise. In *Speech perception and linguistic experience: Issues in cross-language research*, edited by W. Strange. Timonium, MD: York Press.

Jenner, B. (1976). Interlanguage and foreign accent. *Interlanguage Studies Bulletin* 1 (2-3):166-195.

Johansson, S. (1978). *Studies of error gravity: Native reactions to errors produced by Swedish learners of English*. Göteborg, Sweden: Acta Universitatis Gothoburgensis.

Johnson, C.E., and P. Lancaster. (1998). The development of more than one phonology: A case study of a Norwegian-English bilingual child. *International Journal of Bilingualism* 2 (3).

Jongman, A., M. Fourakis, and J.A. Sereno. (1989). The acoustic vowel space of Modern Greek and German. *Language and Speech* 32 (3):221-248.

Joseph, B.D., and I. Philippaki-Warbuton. (1987). *Modern Greek*. London: Croom Helm.

Kellerman, E. (1983). Now you see it, now you don't. In *Language transfer in language learning*, edited by S. Gass and L. Selinker. Rowley, Mass.: Newbury House.

Kondo, Y. (1995). Production of schwa by Japanese speakers of English: a crosslinguistic study of coarticulatory strategies. PhD, Linguistics, University of Edinburgh, Edinburgh.

Koutsoudas, A., and O. Koutsoudas. (1983). A contrastive analysis of the segmental phonemes of Greek and English. In *Second language learning, contrastive analysis, error analysis, and related aspects*, edited by B. Robinett and J. Schachter: University of Michigan Press.

Krashen, S. (1973). Lateralization, language learning, and the critical period: some new evidence. *Language Learning* 23:63-74.

Krashen, S. (1981). *Second language acquisition and second language learning*. Oxford: Pergamon Press.

Ladd, D.R. (1983). Phonological features of intonational peaks. *Language* 59:721-759.

Ladd, D.R. (1988). Declination 'reset' and the hierarchical organization of utterances. *Journal of the Acoustical Society of America* 84:530-44.

Ladd, D.R. (1996). *Intonational phonology*. Cambridge: Cambridge University Press.

Ladd, D.R. (1998). Bruce, Pierrehumbert, and the elements of intonational phonology. In *Festschrift for Gosta Bruce*, edited by M. Home: Kluwer.

Ladd, D.R., A. Arvaniti, and I. Mennen. (1997). Toward a typology of phrase tones. Paper read at Linguistics Association of Great Britain, Spring Meeting, at Edinburgh.

Ladd, D.R., I. Mennen, and A. Schepman. (forthcoming). Phonological conditioning of peak alignment in rising pitch accents in Dutch. .

Ladd, D.R, and J. Terken. (1995). Modelling intra- and inter-speaker pitch range variation. *Proceedings of the thirteenth international congress of phonetic sciences* 2:386-389.

Ladefoged, P. (1993). *A course in phonetics*. 3 ed: Harcourt Brace & Company.

Lado, R. (1957). *Linguistics across cultures*. Ann Arbor, MI: University of Michigan Press.

Leather, J., and A. James. (1996). Second language speech. In *Handbook of second language acquisition*, edited by W. C. Ritchie and T. K. Bhatia. San Diego, London: Academic Press.

Leben, W. (1973). Suprasegmental phonology. PhD, MIT.

Leben, W. (1976). The tones in English intonation. *Linguistic Analysis* 2:69-107.

Lenneberg, E. (1967). *Biological foundations of language*. New York: Wiley.

Lennon, P. (1990). Investigating fluency in EFL: a quantitative approach. *Language Learning* 40:387-417.

Lepetit, D. (1989). Cross-linguistic influence on intonation: French/Japanese and French/English. *Language Learning* 39:397-413.

Lieberman, A.M. (1996). *Speech: a special code*. Edited by L. Gleitman, S. Carey, E. Newport and E. Spelke, *Learning, development, and conceptual change*. Cambridge, Massachusetts: The MIT Press.

Lieberman, M. (1975). The intonational system of English. PhD, MIT.

Lieberman, M., and J. Pierrehumbert. (1984). Intonational invariance under changes in pitch range and length. In *Language Sound Structure*, edited by M. Aronoff and R. Oehrle. Cambridge, MA: MIT Press.

Lim, L. (1995). A contrastive study of the intonation patterns of Chinese, Malay and Indian Singapore English. *Proceedings of the thirteenth international congress of phonetic sciences*:402-405.

Long, M.H. (1990). Maturation constraints on language development. *Studies in Second Language Acquisition* 12 (3):251-285.

Low, E.L., and E. Grabe. (in press). A contrastive study of prosody and lexical stress placement in Singapore English and British English. *Language and Speech* 42 (1).

Mackridge, P. (1985). *The Modern Greek language: A descriptive analysis of Standard Modern Greek*. Oxford: Oxford University Press.

Magen, H. (1998). The perception of foreign-accented speech. *Journal of Phonetics* 26:381-400.

Major, R.C. (1986). Paragoge and degree of foreign accent in Brazilian English. *Second Language Research* 2:53-71.

Major, R.C. (1990). L2 acquisition, L1 loss, and the critical period hypothesis. In *New Sounds 90: Proceedings of the 1990 Amsterdam Symposium on the Acquisition of Second-language Speech*, edited by J. Leather and A. James, Amsterdam: University of Amsterdam.

Malikouti-Drachman, A., and G. Drachman. (1980). Slogan chanting and speech rhythm in Greek. In *Phonologica 1980*, edited by W. Dressler, O. Pfeiffer and J. Rennison. Innsbruck.

Malliopoulos, C., and G. Carayannis. (1997). An F0 generation algorithm for the Greek language. *The ESCA workshop on intonation: Theory, models and applications*:219-222.

Markham, D. (1997). *Phonetic imitation, accent and the learner*. Lund: Lund University Press.

Mennen, I. (1993). Ο νεοελληνικός επιτονισμός: Μια ανάλυση του επιτονισμού σε ερωτηματικές και καταφατικές προτάσεις (Modern Greek Intonation: an analysis of the intonation in questions and statements). Doctoraalscriptie, Department of Modern Greek and Byzantinology, University of Amsterdam, Amsterdam.

Mirambel, A. (1959). *La langue grècque moderne*. Paris: Librairie Klincksieck.

Munro, M., and T. Derwing. (1995). Processing time, accent, and comprehensibility in the perception of native and foreign-accented speech. *Language and Speech* 28:389-406.

Munro, M.J. (1995). Nonsegmental factors in foreign accent: Ratings of filtered speech. *Studies in Second Language Acquisition* 17:17-34.

Neufeld, G. (1977). Language learning ability in adults: A study on the acquisition of prosodic and articulatory features. *Working Papers in Bilingualism* 12:45-60.

Neufeld, G. (1979). Towards a theory of language learning ability. *Language Learning* 29:227-241.

Neufeld, G. (1987). On the acquisition of articulatory and prosodic features in adult language learning. In *Interlanguage phonology: The acquisition of a second language sound system*, edited by G. Ioup and S. H. Weinberger. Cambridge, Mass.: Newbury House.

Nolan, F., and E. Grabe. (work in progress). English intonation in the British isles. Cambridge: University of Cambridge.

Novoa, L., D. Fein, and L. Obler. (1988). Talent in foreign languages: A case study. In *The exceptional brain: Neuropsychology of talent and special abilities*, edited by L. Obler and D. Fein. New York: Guilford.

O'Connor, J.D., and G.F. Arnold. (1973). *Intonation of colloquial English*. London: Longman.

Odlin, T. (1989). *Language transfer*. Edited by M. H. Long and J. C. Richards, *The Cambridge applied linguistics series*. Cambridge: Cambridge University Press.

Oyama, S. (1976). A sensitive period for the acquisition of the non-native phonological system. *Journal of Psycholinguistic Research* 5:261-283.

Palmer, H. (1992). *English intonation, with systematic exercises*. Cambridge: Heffer.

Papazachariou, D. (1997). Language variation and the social construction of identity: The sociolinguistic role of intonation among adolescents in Northern Greece. PhD, Department of Language and Linguistics, University of Essex.

Penfield, W., and L Roberts. (1959). *Speech and brain-mechanisms*. Princeton, NJ: Princeton University Press.

Pennington, M. (1992). Discourse factors related to L2 phonological proficiency: an exploratory study. In *Proceedings of the 1992 Amsterdam Symposium on the acquisition of second-language speech*, edited by J. Leather and A. James. Amsterdam: University of Amsterdam.

Pierrehumbert, J.B. (1980). The phonology and phonetics of English intonation. PhD, MIT, Cambridge, MA.

Pierrehumbert, J.B., and M.E. Beckman. (1988). *Japanese tone structure*. Cambridge, MA: MIT Press.

Polivanov, E. (1931). La perception des sons d'une langue étrangère. *Travaux du Cercle Linguistique de Prague* 4:79-96.

Prieto, P., J. Van Santen, and J. Hirschberg. (1995). Tonal alignment patterns in Spanish. *Journal of Phonetics* 23:429-453.

Raupach, M. (1980). Temporal variables in first and second language speech production. In *Temporal variables in speech: studies in honour of Frieda Goldman-Eisler*, edited by H. Dechert and M. Raupach. The Hague: Mouton.

Rietveld, T., and C. Gussenhoven. (1995). Aligning pitch targets in speech synthesis: effects of syllable structure. *Journal of Phonetics* 23:375-385.

Rochet, B.L. (1995). Perception and production of second language speech sounds by adults. In *Speech perception and linguistic experience: Issues in cross-language research*, edited by W. Strange. Timonium MD: York Press.

Rose, P. (1987). Considerations in the normalisation of the fundamental frequency of linguistic tone. *Speech Communication* 6:343-51.

Rvachew, S., and D.G. Jamieson. (1995). Learning new speech contrasts: Evidence from adults learning a second language and children with speech disorders. In *Speech perception and linguistic experience: Issues in cross-language research*, edited by W. Strange. Timonium MD: York Press.

Sabin, E., E. Clemmer, D. O'Connell, and K. Sabin. (1979). A pausological approach to speech development. In *Of speech and time: temporal speech patterns and interpersonal contexts*, edited by A. W. Seigman and S. Feldstein. Hillsdale, NJ: Erlbaum.

Schneiderman, E., and C. Desmarais. (1988). The talented language learner. Some preliminary findings. *Second language research* 4:91-109.

Scovel, T. (1969). Foreign accents, language acquisition, and cerebral dominance. *Language Learning* 1 (3 & 4):245-253.

Scovel, T. (1981). The recognition of foreign accents in English and its implications of psycholinguistic theories of language acquisition. In *Proceedings of the 5th Congress of AILA*, edited by J.-G. Savard and L. Laforge. Laval, Canada: University of Laval Press.

Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics* 10:209-231.

Selinker, L. (1992). *Rediscovering interlanguage*. Edited by C. N. Candlin. 1 ed, *Applied linguistics and language study*. New York: Longman.

Setatos, M. (1974). *Phonology of Standard Modern Greek (in Greek)*. Athens: Papazisis.

Sharwood Smith, M. (1994). *Second language learning: theoretical foundations, Applied linguistics and language study*. London: Longman.

Sheldon, A., and W. Strange. (1982). The acquisition of /r/ and /l/ by Japanese learners of English: evidence that speech production can precede speech perception. *Applied Psycholinguistics* 3:243-261.

Silverman, K., and J. Pierrehumbert. (1990). The timing of prenuclear high accents in English. In *Papers in Laboratory Phonology*, edited by J. Kingston and M. E. Beckman. Cambridge, MA: Cambridge University Press.

Sorace, A. (1993). Incomplete vs. divergent representations of unaccusativity in non-native grammars of Italian. *Second Language Research* 9 (1):22-47.

Steele, J. (1986). Nuclear accent F0 peak location: effect of rate, vowel and number of syllables. *Journal of the Acoustical Society of America* 80 (Supplement 1):s51.

Stockwell, R., and D.J. Bowen. (1965). *The sounds of English and Spanish*. Chicago: Chicago Press.

Strange, W. (1995a). Cross-language studies of speech perception: a historical review. In *Speech perception and linguistic experience: Issues in cross-language research*, edited by W. Strange. Timonium MD: York Press.

Strange, W. (1995b). Phonetics of second-language acquisition: past, present, future. *Proceedings of the fourteenth international congress of phonetic sciences* 4:76-83.

Strange, W., ed. (1995c). *Speech perception and linguistic experience: Issues in cross-language research*. Timonium MD: York Press.

Tarone, E. (1987). The phonology of interlanguage. In *Interlanguage phonology: The acquisition of a second language sound system*, edited by G. Ioup and S. H. Weinberger. Cambridge, Mass.: Newbury House.

't Hart, J. (1976). Psychoacoustic backgrounds of pitch contour stylisation. *IPO Annual Progress Report* 11:11-19.

't Hart, J. (1981). Differential sensitivity to pitch distance, particularly in speech. *Journal of the Acoustical Society of America* 69 (3):811-821.

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

't Hart, J. , and R. Collier. (1975). Integrating different levels of intonation analysis. *Journal of Phonetics* 3:235-255.

't Hart, J. , R. Collier, and A. Cohen. (1990). *A perceptual study of intonation: and experimental-phonetic approach*. Cambridge: Cambridge University Press.

Trubetzkoy, N. (1939). *Grundzüge der Phonologie*. Berkeley: University of California Press.

Tsakirides, E. (1980). *Spoken Greek: an audio-visual course. Step 1*. Venice, Ca.: Tsakirides.

Van Els, T., and K. de Bot. (1987). The role of intonation in foreign accent. *The Modern Language Journal* 71:147-155.

Van Katwijk, A. (1974). *Accentuation in Dutch: An experimental linguistic study*. Assen: Van Gorcum.

Van Zanten, E., L. Damen, and E. Van Houten. (1991). The ASSP Speech Database: SPIN/ASSP-report.

Verhoeven, J. (1991). Perceptual aspects of Dutch intonation, PhD, University of Edinburgh, Edinburgh.

Verhoeven, J. (1994). The discrimination of pitch movement alignment in Dutch. *Journal of Phonetics* 22:65-85.

Vogel, I. (1991). Prosodic phonology: second language acquisition data as evidence in theoretical phonology. In *Crosscurrents in second language acquisition and linguistic theories*, edited by T. Huebner and C. A. Ferguson. Amsterdam: Benjamins.

Waring, H. (1976). The intonation of Modern Greek. PhD, University of London, London.

Waring, H. (1982). The investigation of Modern Greek intonation. *Mandatophoros* 20:18-29.

Weinberger, S. (1990). Minimal segments in second language phonology. In *New Sounds 90, Proceedings of the 1990 Amsterdam Symposium on the Acquisition of Second-Language Speech*, edited by J. Leather and A. James. Amsterdam: University of Amsterdam.

Weiss, R. (1976). The role of perception in teaching the German vowels to American students. In *Proceedings of the Fourth International Congress of Applied Linguistics*, edited by G. Nickel. Stuttgart: Universität Stuttgart.

Wenk, B. (1985). Speech rhythms and second language acquisition. *Language and Speech* 28:157-175.

Wenk, B. (1986). Crosslinguistic influence in second language phonology: speech rhythms. In *Cross-linguistic influence in second language acquisition*, edited by E. Kellerman and M. Sharwood Smith. Oxford: Pergamon.

Willems, N. (1982). *English intonation from a Dutch point of view*. Dordrecht: Foris Publications.

Williams, L. (1979). The modification of speech perception and production in second language learning. *Perception and Psychophysics* 21:289-297.

Williams, L. (1980). Phonetic variation as a function of second-language learning. In *Child phonology: Perception*, edited by G. Yeni-Komshian, J. Kavanagh and C. Ferguson. New York: Academic Press.

Wode, H. (1976). Developmental sequences in naturalistic second language acquisition. *Working Papers in Bilingualism* 11:1-31.

Wode, H. (1980). Phonology in L2 acquisition. In *Second language development: Trends and issues*, edited by S. Felix. Tübingen: Gunter Narr.

Yamada, R.A., Y. Tohkura, A.R. Bradlow, and D.B. Pisoni. (1995). Effects of perceptual learning on the production of a non-native contrast. Paper read at Journal of the Acoustical Society of America.

Zobl, H. (1984). Cross-language generalizations and the contrastive dimension of the interlanguage hypothesis. In *Interlanguage*, edited by E. Davies, C. Cripser and A. Howatt. Edinburgh: Edinburgh University Press.

## Appendix A: Materials for Dutch prenuclear accents

There are two sets of materials, one with phonologically short vowels (set DG-short: sentences 1 to 20) and one with phonologically long vowels (set DG-long: sentences 21 to 40). The target words are underlined, and transcribed phonemically at the end of each sentence. Although the phonologically long high vowels are usually indicated by a 'half-length' mark in Dutch, all phonologically long vowels are indicated with the standard IPA length mark.

Set D-short:

1. Hij wilde de rillende kinderen tracteren op warme chocolademelk. /'rɪləndə/  
He wanted to buy the shivering children some hot chocolate.
2. Met haar beminnelijk gedrag kon ze iedereen om haar vinger winden. /bə'mɪnələk/  
With her amiable behaviour she could wrap everyone around her little finger.
3. Je moet haar bedillerig gepraat maar langs je heen laten gaan. / bə'dɪlərəx/  
You should just ignore that meddling talk.
4. Hij kon van de Limburgse vlaaien maar geen genoeg krijgen. /'lɪmbʏrksə/  
He could not get enough of the Limburg fruit tarts.
5. Wij konden de rennende atleten met geen mogelijkheid bijhouden. /'rənəndə/  
There was no way we could keep up with the running athletes.
6. Hij kon de remmende auto nog net ontwijken. /'rɛməndə/  
He only just managed to avoid the braking car.
7. Ze wilden hun belemmerend schoeisel zo gauw mogelijk uittrekken. / bə'lemərənt/  
They wanted to remove their constricting footwear as soon as possible.

8. Zij had de bedremmelde meisjes beter naar huis kunnen sturen. / bə'dreməldə/  
It would have been better if she had sent the embarrassed girls home.
9. Ze waren door hun mollige gezichten ongeschikt voor modellenwerk. /'mɒləxə/  
Because of their plump faces, they were unsuitable for modelling.
10. Ik kon mijn morrende collega niet meer luchten of zien. /'mɔrəndə/  
I could no longer stand the sight of my grumbling colleague.
11. Hij kon met zijn rollende ogen alle kinderen angst aanjagen. /'rɒləndə/  
He could frighten all the children with his rolling eyes.
12. We kunnen de mondige studenten van tegenwoordig niet meer de baas.  
/'mɒndəxə/  
We can't keep the upper hand with today's assertive students.
13. Ze hebben de genummerde kwitanties helemaal door elkaar gegooid.  
/xə'nɪmərðə/  
They have messed up the order of the numbered invoices.
14. Je kan de lullige verhalen van Willem beter niet serieus nemen. /'lɪləxə/  
You shouldn't take Willem's cruddy stories seriously.
15. Hij was door zijn lummelige houding niet bepaald aantrekkelijk. /'lɪmələxə/  
His oafish attitude did not make him particularly attractive.
16. Hij zag de brullende gorilla's opgewonden heen en weer rennen. /'brɪləndə/  
He saw the roaring gorillas run up and down full of excitement.
17. Ze wordt door haar mannelijk karakter vaak voor een man versleten. /'mænələk/  
Because of her manly behaviour, she is often mistaken for a man.
18. Hij kan zijn rammelend betoog maar beter herschrijven. /'ræmələnt/  
He'd better rewrite his rambling story.
19. Je kon aan zijn lallende uitspraak gelijk horen dat hij dronken was. /'læləndə/  
You could tell he was drunk from his slurred pronunciation.

20. Ik was de drammende telefoontjes van dat vervelende mens spuugzat.  
/ˈdrʌməndə/

I was sick of the pestering phone calls of that annoying woman.

Set D-long:

21. Je moet in aluminium pannen geen azijn gebruiken. /ˌɒlyˈmiːniʊm/

You should not put vinegar in aluminium pans.

22. De jury kon de lyrische stijl van zijn roman niet bepaald waarderen. /ˈliːrɪsə/

The jury did not particularly appreciate his novel's lyrical style.

23. Hij had de alinea's van zijn werkstuk veel te lang gemaakt. /ɑːˈliːniːjəːs/

He had made the paragraphs of his essay far too long.

24. Je moet je dierlijke instinct niet altijd onderdrukken. /ˈdiːrləkə/

You should not always suppress your animal instincts.

25. Ze moesten de lenige assistente uit haar benarde positie bevrijden. /ˈleːnəxə/

They had to free the supple assistant from her awkward position.

26. Ik kan de melige grappen van Seth Gaaikema niet meer aanhoren. /ˈmeːləxə/

I can no longer stand the corny jokes of Seth Gaaikema.

27. Hij heeft voor de dinerende gasten een optreden verzorgd. /diːˈneːrəndə/

He has arranged a performance for the dining guests.

28. Door zijn belerende gepreek joeg hij iedereen de kerk uit. /bəˈleːrəndə/

He managed to empty the church with his pedantic sermons.

29. Je moet bij meerdere geschiktheid je persoonlijke voorkeur laten varen.

/ˈmeːrdərə/

In case of superior suitability, you should abandon your personal preference.

30. Je moet in romige sauzen een scheutje cognac doen; dat is lekker. /ˈrɔːməxə/

You should put some brandy in creamy sauces; that's nice.

31. Hij moest de beloningen van de politie gaan inventariseren. /bə'lo:nɪŋə/  
He had to go and catalogue the police rewards.
32. Het kan in een noordelijk klimaat 's winters flink koud zijn. /'no:rdələk/  
It can get very cold in winter in a Northern climate.
33. Hij wil met de naburige kerkdorpen een belangengroep oprichten. /na:'by:rəxə/  
He wants to start an interest group with the neighbouring villages.
34. Hij kon de ongedurige paarden niet meer in bedwang houden. /ɔnxədy:rəxə/  
He could no longer keep the restless horses under control.
35. Ze willen de Lunense heide weer openstellen voor het publiek. /'ly:nənsə/  
They want to re-open the Lunen Heath to the public.
36. Je moet de duurdere produkten vooraan in de vitrinekast zetten. /'dy:rðərə/  
You should put the more expensive items at the front of the display.
37. We zoeken een dynamische dertiger voor deze veelzijdige functie. /di:'na:misə/  
We are seeking a dynamic person in their thirties for this multi-faceted job.
38. Hij kon de malende gedachten niet uit het hoofd zetten. /'ma:ləndə/  
He could not get the persistent thoughts out of his mind.
39. Hij is in zijn manische periode; dan is hij altijd zo druk. /'ma:nisə/  
He is in his manic phase; he is always active then.
40. Hij moest de dralende studenten tot drie keer toe roepen.  
He had to call the lingering students three times. /'dra:ləndə/

## Appendix B: Materials for Greek prenuclear accents

Set G:

Materials are presented in Greek alphabet first, followed by a broad transcription and a gloss. The target words are underlined.

1. Το κελάρισμα της πηγής ακουγόταν καθαρά πίσω από τους θάμνους.  
[to ce'larizma tis pi'jis aku'gotan kaθa'ra 'piso a'po tus 'θamnus]  
The gurgling of the spring was clearly heard behind the bushes.
2. Η παράδοση των επίπλων θα γίνει την Τρίτη το πρωί.  
[i para'dosi ton e'riplon θa 'jini ti 'driti to pro'i]  
The furniture delivery will take place on Tuesday morning.
3. Το κουράρισμα που του κάνει ο καινούριος γιατρός φαίνεται να έχει αποτελέσματα.  
[to kou'rarizma pu tu 'kani o ce'nurjios ja'tros 'fenete na 'eçi apote'lezmata]  
The cure that the new doctor follows appears to have results.
4. Τα γλυκολέμονα που αγόρασες δεν είναι φρέσκα.  
[ta γliko'lemona pu a'gorases den 'ine 'freska]  
The limes you bought are not fresh.
5. Η ανέραστη θεία τους τους εξέπληξε όλους με τον ξαφνικό γάμο της.  
[i a'nerasti 'θia tus tus ekseplikse 'olus me ton gzafni'ko 'γamo tis]  
Their "loveless" aunt surprised them all with her sudden marriage.

6. Παρά την τουριστική ανάπτυξή τους, τα Ιόνια παραμένουνε τα ομορφότερα νησιά της Ελλάδας.

[pa'ra tin duristi'ci a'naptiksi tus // ta i'onia para'menune ta omorfotera ni'sja tis e'laðas]

Despite tourist development, the Ionian islands remain the most beautiful ones of Greece.

7. Ο βελούδινος καναπές είναι κληρονομιά απ'τη γιαγιά μου.

[o ve'ludinos kana'pes 'ine klirono'mia apti ja'ja mu]

I inherited the velvet sofa from my grandmother.

8. Η μεταβίβαση των τίτλων θα γίνει αύριο στο γραφείο του δικηγόρου.

[i meta'vivasi ton 'titlon ða 'jini 'avrio sto gra'fio tu ðici'goru]

The deeds' transfer will take place tomorrow at the lawyer's office.

9. Όταν επιβραδύναμε το βήμα μας, τα παιδιά μας έφτασαν χωρίς δυσκολία.

[otan epivraðiname to 'vima mas ta peðja mas 'eftasan xo'ris ðisko'lia]

When we slowed down (our step) the children reached us without difficulty.

10. Με το κυνήγημα που του έκανες δεν μου κάνει εντύπωση που δεν σου ξαναμίλησε.

[me to ci'ni'jima pu tu 'ekanes 'ðe mu 'kani e'diposi pu ðen su ksana'milise]

The way you treated him, I'm not surprised he hasn't talked to you again.

11. Το ποδήλατο του Πέτρου κλάπηκε χτες το βράδυ, έξω από το σπίτι του.

[to poðilato tu 'petru 'klapice xtes 'vraði 'ekso a'po to 'spiti tu]

Petros's bike was stolen last night outside his house.

12. Το νυχτολούλουδο που είχαμε στον κήπο μας ξεράθηκε ξαφνικά πέρυσι το καλοκαίρι.  
 [to nixto'luluðo pu 'ixame sto 'jipo mas kse'raθice ksafni'ka 'perisi to kalo'ceri]  
 The "nightflower" we had in our garden died suddenly last summer.
13. Τα ολόλευκα τριαντάφυλλα είναι το αγαπημένο της λουλούδι.  
 [ta o'lolefka tria'ndafila 'ine to aγapi'meno tis lu'ludi]  
All-white roses are her favourite flower.
14. Η ανόρεχτη έκφρασή της μου έκοψε κάθε διάθεση για συζήτηση.  
 [i a'norexti 'ekfra'si tis mu 'ekopse 'kaθe 'diaθesi ja si'zitisi]  
 Her bored expression stops every desire for conversation.
15. Το φιλοδώρημα που του έδωσες δεν ήταν αρκετό κατά τη γνώμη μου.  
 [to filoðorima pu tu 'eðoses ðen 'itan arce'to ka'ta ti 'gnomi mu]  
 The tip you gave him was not enough in my opinion.
16. Τα αναβράζοντα δισκία βιταμίνης C είναι από τις αγαπημένες μου λιχουδιές.  
 [ta ana'vrazonda ði'scia vita'minis ci 'ine a'po tis aγapi'menezmu lixuðjes]  
 The dissolvable tablets of vitamin C are one of my favourite delicacies.
17. Η παρέμβαση του υπουργού δεν έφερε το ποθητό αποτέλεσμα.  
 [i pa'remvasi tu ipur'γy ðen 'efere to poθi'to apo'telezma]  
 The minister's intervention did not have the desired result.

18. Η ανάρμοστη συμπεριφορά της είχε ως αποτέλεσμα την αποβολή της από το σχολείο.  
 [i a'narmosti siberifora tis 'ice os apotelezma tin apovolitis a'po to sxolio]  
 Her improper behaviour resulted in her expulsion from school.
19. Η κυβέρνηση συνασπισμού δεν έμεινε στην ηγεσία πάνω από μερικούς μήνες.  
 [i ci'vernisi sinaspizmu den 'emine stin ije'sia 'pano a'po merikus 'mines]  
 The coalition government did not stay in power for more than a few months.
20. Τα μεταλλεύματα του Παγγαίου όρους ήταν η βασική πηγή εισοδήματος της Αρχαίας Μακεδονίας.  
 [ta meta'levmata tu pa'jeu 'orus 'itan i vasi'ci pi'ji iso'dimatatos tis ar'ceas macedonias]  
 The minerals of Mount Pangaio were the basic source of income of Ancient Macedonia.

## Appendix C: Materials for Greek nucleus-non-final (NNF) yes/no questions

Test materials are presented in IPA transcription, together with a gloss. Only the yes/no questions are transcribed. For an example of the dialogues used the reader is referred to section 2.2 of Chapter 4. The yes/no questions are all expected to have the nuclear accent on the first content word (i.e. they are all nucleus-non-final yes/no questions). The sentences are divided into three sets of 20 sentences each: one with lexical stress on the utterance-final syllable, one with lexical stress on the penultimate, and one with lexical stress on the antepenultimate syllable of the final content word. The test word (i.e. the word bearing the final rise-fall) is underlined.

Final word stress:

1. [ta 'tros ta laðe'ra] "Do you eat food cooked with oil?"
2. [θa bo'rusa na se ðo] "Could I see you?"
3. [na 'katsume eðo] "Shall we sit here?"
4. [sa'resi to me'nu] "Do you like the menu?"
5. [ta 'pires ta ja'la] "Did you buy the glasses?"
6. [ti 'vlepis ti ma'ja] "Have you seen the yeast?"
7. [ðe 'vazis ce ana'na] "Won't you add some pineapple?"
8. [θa 'fas ka'rame'le] "Will you have some crème caramel?"

9. [to 'kseris to yuði] "Do you know Goudi?"
10. [mu ðinis mpa yu'la] "Can you give me a sip?"
11. [ðo'cimases me'ze] "Have you tried the starters?"
12. [to 'ides to mo'ro] "Did you see the baby?"
13. [θa su 'ftasi to ma'li] "Do you have enough wool?"
14. [ksanama'yirepses ra'yu] "Have you made ragout before?"
15. [sas 'eftase i bo'ja] "Was there enough paint?"
16. [ka'θarise i vro'mpa] "Did you manage to remove the stain?"
17. [θa 'valete mpa'lo] "Will you ever learn a lesson?"
18. [ti 'vlepis ti mo'ni] "Can you see the monastery?"
19. [ma'zevi to li'no] "Does linen shrink?"
20. [ti bo'lasan ti mi'la] "Did they graft the apple tree?"

Penultimate word stress:

1. [ðen 'pame sto ka'vuri] "Shall we go to Kavouri?"
2. [a'yorases le'mona] "Did you buy lemons?"
3. [θi'mase pos ti 'lene] "Do you remember her name?"
4. [ti çe'retises ti 'nina] "Did you greet Nina?"

5. [ma xo'rane sti ba'pera] "But do they fit in the bathtub?"
6. [to 'lavate to 'ðema] "Did you get the package?"
7. [se vo'levi mesi'meri] "Does the afternoon suit you?"
8. [ta 'iðes ta yla'roŋa] "Did you see the sea-gulls?"
9. [ti yn'orizis ti ma'rina] "Have you met Marina?"
10. [su a'resi to 'vazo] "Do you like the vase?"
11. [tu 'arese to 'ðoro] "Did he like the present?"
12. [ta ka'θarises ta 'mila] "Did you peel the apples?"
13. [na 'vyume sti veranda] "Shall we go out on the verandah?"
14. [kselabi'karise to nero] "Is the water clean again?"
15. [ma mas xo'rai to divani] "Is there enough room (for us) on the couch?"
16. [ta fu'skosan ta ba'loŋa] "Did they blow up the balloons?"
17. [me to poðilato pi'yeni] "Does he go on the bicycle?"
18. [psi'θikan ta la'zana] "Is the lasagne done?"
19. [le'çazi to me'lani] "Does ink stain?"
20. [to 'θes to maksilari] "Do you need the pillow?"

Antepenultimate word stress:

1. [ton ep'iskevase to ne'romilo] "Did they repair the watermill?"
2. [a'vɣa su kaθa'rizune] "What's so funny about it?"
3. [tin ana'ɣnorisate tin 'iriða] "Did you recognise Irida?"
4. [ksana'piɣate stin 'eyina] "Have you visited Aegina before?"
5. [tin ɣnorizete tin 'elena] "Have you met Helena?"
6. [su a'resun ta ks'inomila] "Do you like crab-apples?"
7. [ton 'ides ton xa'riðimo] "Have you seen Charidimos?"
8. [pe'risepse ri'zoɣalo] "Is there any rice-pudding left?"
9. [a'liθja to no'mizune] "Do they really think that?"
10. [θa 'fiɣun apto 'livano] "Will they leave Lebanon?"
11. [a'ɣi to laðo'lemono] "Does the oil-and-lemon sauce take long?"
12. [ma ori'masan ta ko'romila] "But are the wild plums ripe?"
13. [ma 'erikse ço'nonero] But was it raining sleet?"
14. [θa 'ftasi to roðonero] "Is there enough rosewater?"
15. [θa ton pro'lavume ton 'virona] "Will we be in time for Byron?"
16. [to θi'mase to 'numero] "Do you remember the number?"

17. [θa me si'kosi to mɔ'noziyo] "Do you think the horizontal bar will bear me?"

18. [ta'ɣorasan ta o'moloya] "Did they buy the bonds?"

19. [pame tourko'limano] "Shall we go to the Turkish harbour?"

20. [si'niyoro se 'valane] "Did they appoint you as their council?"

## Appendix D. Normalisation method

This appendix presents the normalisation method which was used in this thesis, that is a method used to abstract away from differences in pitch range between speakers. In chapter 1 (section 5.3.2) it was noted that one of the difficulties in comparing intonation across languages is that pitch shows a great deal of inter- and intra-speaker variation. These differences between and within speakers may obscure general patterns of intonation. In Figure D.1 (reprinted from figure 4.1) an example is shown of such between-speaker differences in pitch. Although the pattern looks fairly similar for some of the speakers, it can be seen that there are considerable differences in the pitch range of the speakers. Earle (1975) in his study of lexical tones in Vietnamese, developed a normalising model which attempts to abstract away from such between-speaker differences. In this model, F0 values were expressed on a speaker-specific normalised scale which was derived by assigning a value of 100 to the top and a value of 0 to the bottom of the speakers' overall F0 range. Earle's aim was "to make the tone systems for different speakers more comparable to each other by normalizing for different speaker's characteristic fundamental frequency levels and ranges" (p.113). His assumption was that when the values obtained for each lexical tone are plotted on a normalised scale, the differences between the speakers should be ironed out and the characteristics of each tone should be the same for all speakers. Thus, if this approach were applied to the data in Figure D.1, the differences between the speakers at the various measurement points would have disappeared and all measurement values would come out as virtually the same.

A similar approach to normalising F0 data was taken by Rose (1987), in his study of Wu Chinese. However, his approach differs from Earle's in the sense that he normalised according to the overall average and the units of standard deviation each speaker is removed from the mean.

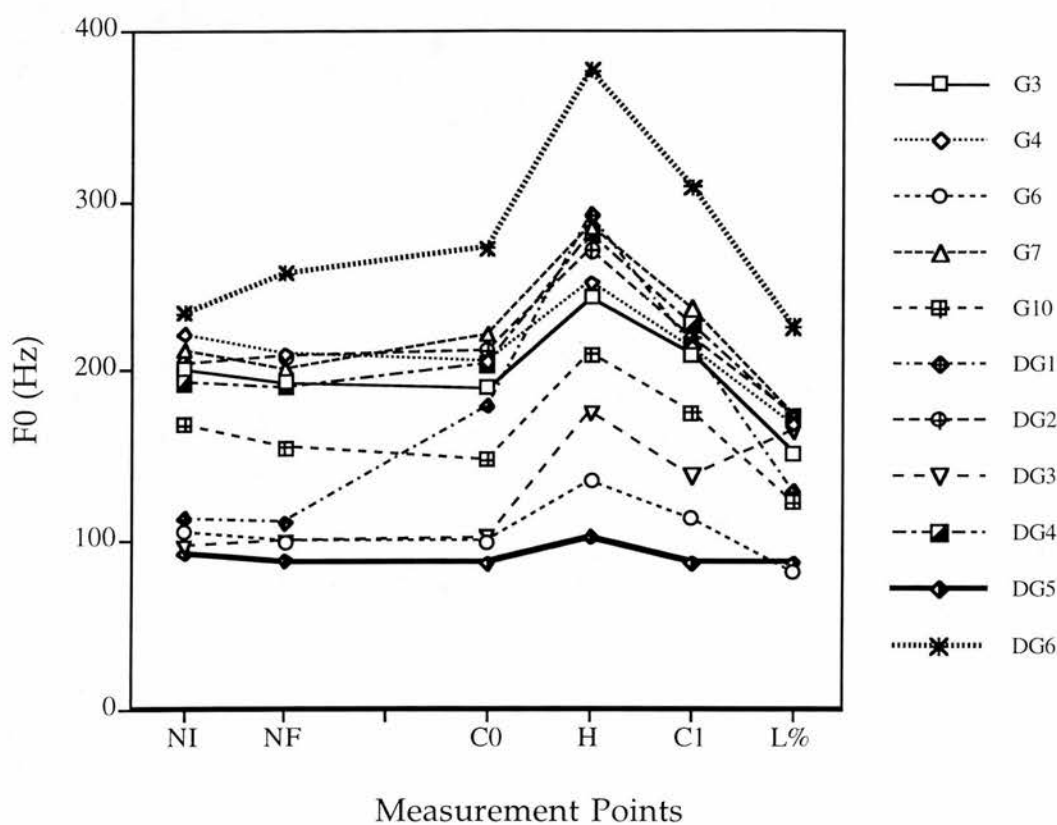


Figure D.1. Raw F0 values for each of the speakers of Group G and DG on various measurement points in Greek NNF yes/no questions. The measurement points are: the onset (NI) and offset (NF) of the accented vowel of the first content word; the consonant onset of the stressed syllable of the final content word (C0); the highest F0 in the final rise-fall (H); the offset of the stressed syllable of the final content word (C1); and the utterance-final low F0 (L%).

Although Rose argues that his approach to normalisation is superior to Earle's, it is clear that the search for normalisation is not over, and that there may be more sophisticated choices for defining the normalised scale than the ones attempted so far. However, both Rose's and Earle's attempts make clear that normalisation is feasible, as their results show a high degree of between-speaker agreement in their description of lexical tones.

In Chapter 4 it was mentioned that inspection of the yes/no question data of Group G and Group DG suggested that there was a difference in the height of the final peak between native and non-native speakers. However, because of differences in the pitch range of individual speakers, it was impossible to decide on the basis of raw F0 data, whether these differences are attributable to group differences, or are due to individual speakers' differences in pitch range. As the aim of this study is to compare the pitch scaling between the group of native and non-native speakers, it is crucial to be able to decide whether the differences found between the two groups of speakers are indeed attributable to group differences. For this reason, it was thought necessary to abstract away from these differences between speakers. Therefore, it was decided to follow the approach taken by Earle (1975) and express raw F0 values on a speaker-specific normalised scale which was derived by assigning a value of 100 to the top and a value of 0 to the bottom of the speakers' overall F0 range.

However, following Earle's approach proved not that easy. First of all, it was impossible to assign a value of 100% to the peak of the yes/no questions. If we had done so, we would have made the peak equal for both groups, and it was exactly the height of the peak which was thought to be different for the two groups. Therefore, in order to capture this difference, another point needed to be found to define the top of the range. The measure which was eventually settled on for defining the top of the range will be described later in this appendix.

A second problem was to do with the measurement unit which would be used to define the speaker-specific normalised scale. That is, would it be better to express the raw values which will be used to define the

speaker-specific normalisation scale in Hz, or in ST or ERB? Earle (1975) expressed the F0 values in Hz. It was decided here to express the F0 values in ERB, before proceeding with the speaker-specific normalisation. There is some evidence (Hermes and Van Gestel, 1991; Hermes and Rump, 1994; Ladd and Terken, 1995) that the ERB scale factors out some of the differences between speakers (specifically differences between female and male voices). It was thought that by expressing the raw values in ERB, before defining a speaker-specific normalisation scale could make normalisation more successful. Therefore, in the experiments described in Chapter 4, the F0 (Hz) values were converted to ERB values, according to the following formula (Rietveld and Van Heuven, 1997: p. 369):

$$ERB = 16.6 \times \log \left( 1 + \frac{F0}{165.4} \right)$$

As mentioned before, impressionistically it seemed that there were differences in the height of the peak between Group D and Group DG. Therefore, it was not possible to assign a value of 100% to the peak of the yes/no questions, as this would make the peak equal for the two groups and the differences between the two groups would disappear. As a consequence it was necessary to find another measure to define the top of the speakers' range. Therefore, data of different sentence types needed to be used in order to find an appropriate measure for the top and the bottom of the speakers' range.

It was thought that the peak and the utterance-final low value (in ERB) of wh-questions would be appropriate measures for defining the top and bottom respectively of each speaker's range. Greek wh-question intonation almost invariably has a sharp rise on the initial question word, followed by a steep fall and (if the question is long enough) a low level stretch. At the very end of the sentence there is often, but not always, a rise (Arvaniti and Ladd, forthcoming). This pattern is illustrated in Figure D.2, where it is presented together with a Greek yes/no

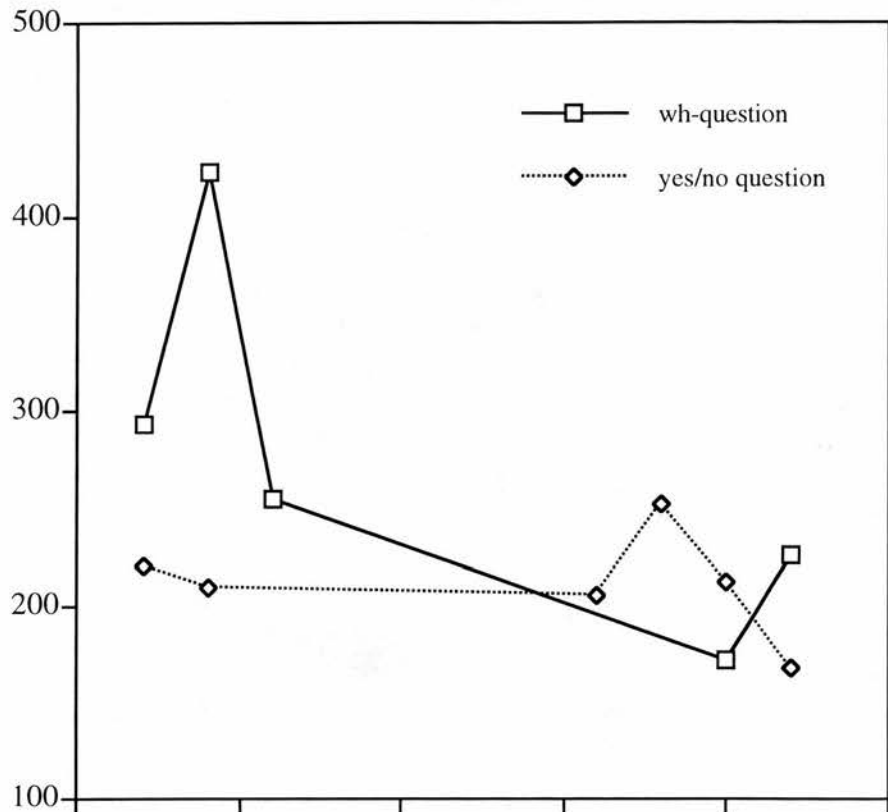


Figure D.2. Illustration of a Greek *wh*-question and nucleus-non-final *yes/no* question. The peak of the *wh*-question which is used to define the top of a speaker's range is quite high compared to the peak in the *yes/no* question. The low value near the end of the *wh*-question (which is used to define the bottom of a speaker's range) is close to the final low in the *yes/no* question.

question. From this figure, it can be seen that the peak of the wh-question (which is used to define the top of a speaker's range) is quite high compared to the peak in the yes/no question. The low value near the end of the wh-question (which is used to define the bottom of a speaker's range) is close to the final low in the yes/no question.

After it was decided to use the initial peak and the final low of wh-questions to define the normalisation scale, the yes/no questions of the speakers of Group G were expressed on this scale. Once it was shown that the speakers of this group showed a high degree of agreement, the same normalising method could be used for Group DG, and the normalised data of the two groups could be compared.

However, there was a further complication in the sense that we did not have the same data for all the speakers. For the speakers of Group G, there were only wh-question data available for speakers G4, G6, and G10. For each of these three Greek speakers data available for the mean value (in ERB) of the initial peak and the utterance-final low, were averaged over 37 wh-questions. Then, for each speaker a value of 100% was assigned to their mean value of the peak, and a value of 0% to their average utterance-final low value in the wh-questions. With the top and bottom of the three speaker's ERB range thus specified, their ERB values on specific measurement points in the yes/no questions were expressed relative to these points. The measurement points were points in the utterance which were thought (after visual inspection of the data) to capture the characteristics of the Greek nucleus-non-final yes/no question contour. The measurement points are: the onset (NI) and offset (NF) of the accented vowel of the first content word; the onset of the stressed syllable of the final content word (C0); the highest F0 in the final rise-fall (H); the offset of the stressed syllable of the final content word (C1); and the utterance-final low F0 (L%).

Thus, each of the three speaker's individual ranges is normalised according to the following formula:

$$NORMerb = \frac{ERB - \bar{L}_i}{\bar{H}_i - \bar{L}_i} \times 100$$

where ERB is any fundamental frequency value for a given speaker expressed in ERBs; NORMerb ('normalised ERB') expresses ERB on a percentage scale of the individual speaker's pitch range; and average  $L_i$  and average  $H_i$  are an individual speaker's average utterance-final low and peak value respectively. Thus, the normalised values for speaker G4 (average  $L_i = 5.21$  and average  $H_i = 14.51$ ) at the final peak are  $((6.67 - 5.21)/9.3) \times 100 = 15.70\%$ .

After the normalised values for these three speakers were established, two points in the yes/no question contour were chosen, one at the initial low stretch before the final rise-fall, and one at the final rise-fall. These points were the NI (the onset of the accented vowel of the first content word) and the H peak (i.e. the utterance-final peak), respectively. The values obtained for these two points were averaged over the three speakers, and used to define the average values of the other speakers (G3 and G7, for whom there were no wh-question data available) at these two points on a normalised scale. Thus, speaker G3 and G7's individual mean ERB values at the H and at NI were set to the other speaker's average normalised values at these points, and their ERB values for all the measurement points were defined relative to these points. This procedure was followed for both the NNF and the NF yes/no questions.

The effectiveness of the scaling can be seen in Figure D3, which shows each speaker's normalised values plotted for each of the measurement points in NNF yes/no questions. All the values now cluster rather tightly, and differences between speakers, such as differences caused by differences in sex, have been highly reduced. The only points where the values do not cluster that tightly are C0 (i.e. the onset of the stressed syllable of the final content word) and L% (the utterance-final low), which are slightly higher for speaker G7. In fact, speaker G7 is the only speaker of Group G who has not been raised in Athens and does not speak standard Athenian Greek. Apparently, her Peloponnesian accent differs from that

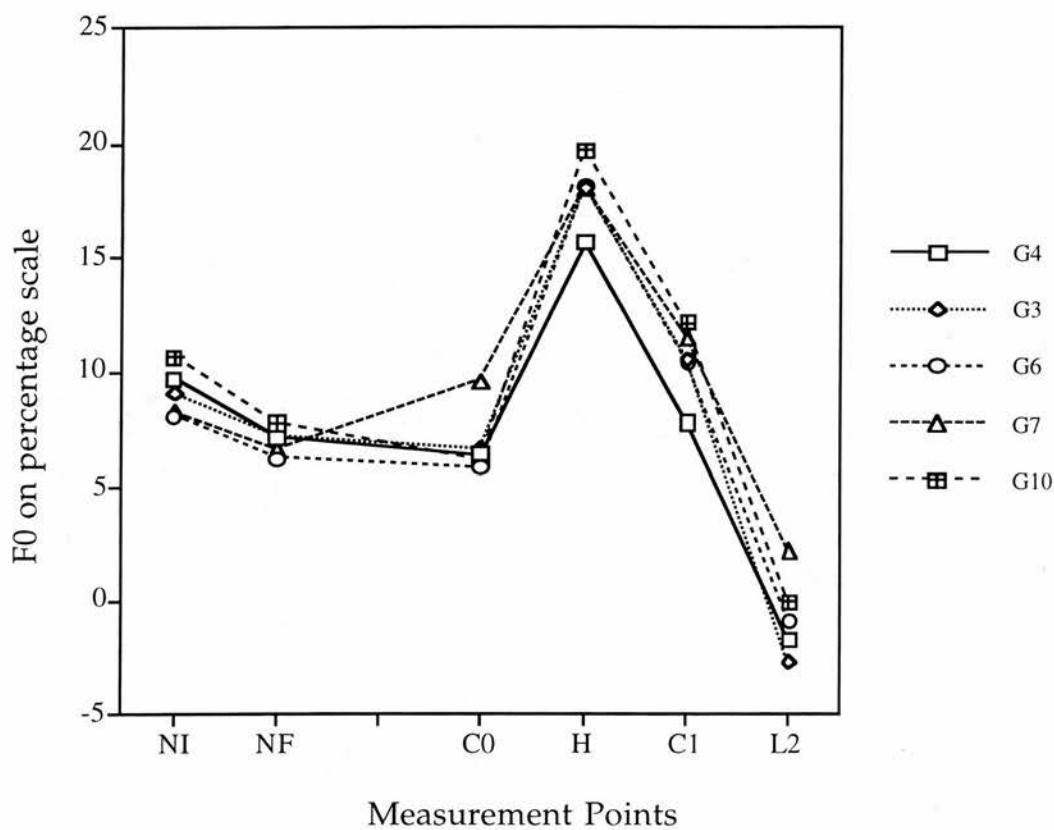


Figure D.3. Normalised values (expressed on a speaker-specific or percentage scale) at the different measurement points in nucleus-non-final yes/no questions for each of the speakers of Group G.

of the other speakers not only at the segmental level (as was thought at the time of recording), but also affects her intonational system slightly.

In order to compare the normalised data of Group G with those of Group DG, ideally the normalisation of group DG should be based on the same sentence type as for Group G. However, there were no wh-questions available for the speakers of Group DG. Therefore, normalisation had to be based on their Greek statements (i.e. the prenuclear accents from Experiment 2).

For this reason, for speakers G3 and G4 (i.e. the only speakers of Group G for which prenuclear accent data were available), normalised values were calculated for their prenuclear accent by using the same scale as previously used for their yes/no questions. In this way, mean percentage values were obtained for their prenuclear L and H tones. These mean percentage values were then used to define the individual scales for each speaker of Group DG. If, for example, the mean percentage value for speakers G3 and G4 of the L is 9%, and that for the H is 25%, then for each speaker of Group DG, the ERB values obtained for the L and H tones of the prenuclear accents are set to these percentages. In this way, individual speakers' scales are established, and their values obtained for the different measurement points of their yes/no questions can be expressed relative to these points.

Although, admittedly, not the most direct way of deriving a normalised scale, it seems effective enough for the purposes of this study. Figure D.4 shows the results of the speakers of both groups expressed on a normalised scale. It can be seen that even though the differences between the speakers are reduced considerably (when compared to the raw F0 data shown in Figure D.1), there still remain differences between the two groups of speakers.

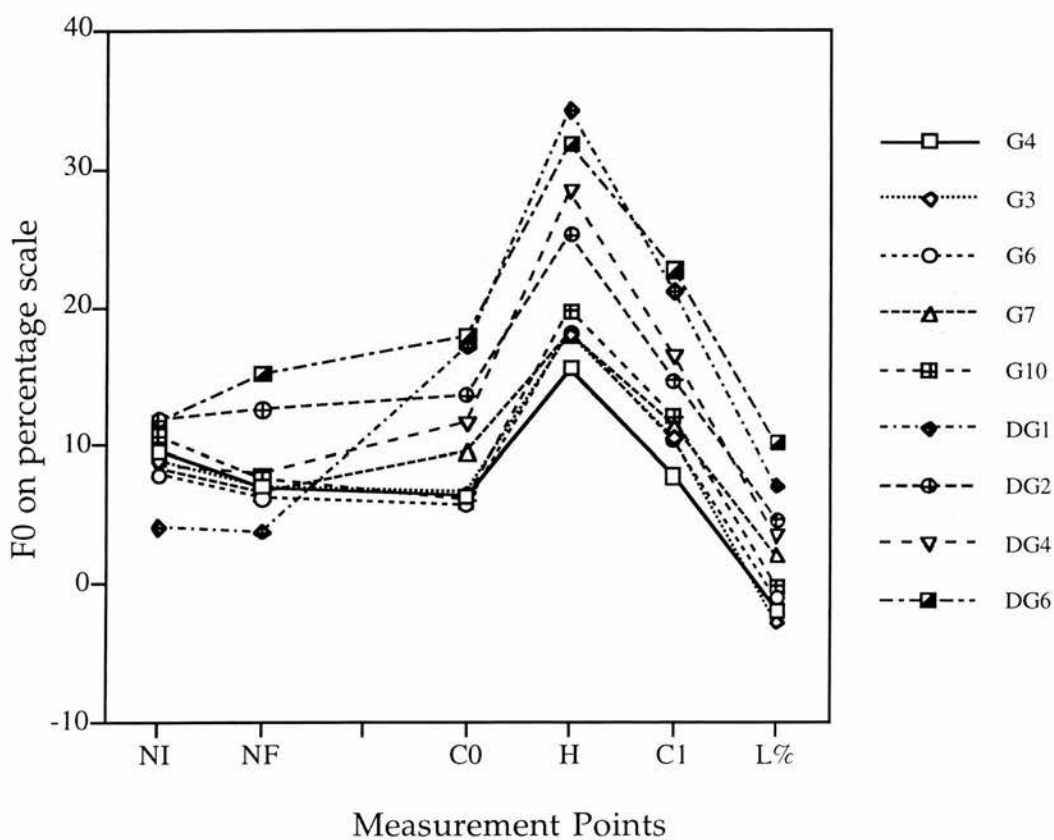


Figure D.4. Normalised values (expressed on a speaker-specific or percentage scale) at the different measurement points for each of the speakers of Group G and Group DG.

## Appendix E: Materials for Greek nucleus-final (NF) yes/no questions

Test materials are presented in IPA transcription, together with a gloss. Only the yes/no questions are transcribed. For an example of the dialogues used the reader is referred to section 2.2 of Chapter 4. The yes/no questions are all expected to have the nuclear accent on the first content word (i.e. they are all nucleus-final yes/no questions). The sentences are divided into three sets of 20 sentences each: one with lexical stress on the utterance-final syllable, one with lexical stress on the penultimate, and one with lexical stress on the antepenultimate syllable of the final content word. The test word (i.e. the word bearing the final rise-fall) is underlined.

Final word stress:

1. [θa 'pane sto ba'li] "Are they going to Bali?"
2. [pire to proi'no] "Did he take the morning train?"
3. [mi'lai sova'ra] "Is he serious?"
4. [e'çi maيدا'no] "Did you use parsley?"
5. [afti me to be're] "The one with the baret?"
6. [ine vor'i'no] "Is it in the north?"
7. [ine ya'la'na] "Are they blue?"

8. [f'iyane ma'zi] "Did they leave together?"
9. [evale jia'la] "Did he get glasses fitted?"
10. [perpa'tusan anga'la] "Did they walk arm in arm?"
11. [exo mela'na] "Do I have a bruise?"
12. [e'ðo ariste'ra] "Here to the left?"
13. [su 'fenete almiro] "Does it appear salty?"
14. [ekane zi'mpa] "Did he make a mess?"
15. [itane ka'lo] "Was it good?"
16. [e'çi gale'ri] "Does he run a gallery?"
17. [na 'ine jerani] "Do you reckon they are cranes?"
18. [e'çi randevu] "Does she have a date?"
19. [kapjo maya'zi] "A shop?"
20. [mi'la me tin iro] "Is she speaking to Iro?"

Penultimate word stress:

1. [foraye ti 'yuna] "Was she wearing her fur coat?"
2. [isuna sto 'bapo] "Were you in the bath?"
3. [tavale sti 'jala] "Did she put them in the bowl?"

4. [xti'pai to 'kuḑuni] "Is the bell ringing?"
5. [ine θimomeni] "Is she angry?"
6. [f'oruse ʧe ple'reza] "Was she also wearing a mourning veil?"
7. [tavale ma'zimu] "Has she taken against me?"
8. [meta'komisan sto ḑoma] "Did they move into the apartment?"
9. [pinis lemo'naḑa] "Are you drinking lemonade?"
10. [ɣnorises ti 'rena] "Did you know Rena?"
11. [sta 'mesa tu ye'nari] "In the middle of January?"
12. [se 'pirakse i 'jiri] "Did the pollen affect you?"
13. [eʧi lano'lini] "Does it contain lanoline?"
14. [vazete le'vanda] "Do you use lavender?"
15. [anevice sto 'vima] "Did he go on the forum?"
16. [θa se peri'meni i mari'lena] "Will Marilena wait for you?"
17. [psaxnis ja ve'lona] "Are you looking for a needle?"
18. [iḑes to ma'jomu] "Have you seen my swimsuit?"
19. [iḑes to sti'lomu] "Have you seen my pen?"
20. [menune sti 'romi] "Do they live in Rome?"

Antepenultimate word stress:

1. [leʝane vromoloʝa] "Were they using dirty words?"
2. [kanis avɣo'lemono] "Are you making egg-and-lemon sauce?"
3. [pa'tuses sta 'vromonera] "Did you step in the sewage?"
4. [fo'resate ʝe 'malina] "Were you even wearing woollies?"
5. ['vazis ðendro'livano] "Do you use rosemary?"
6. [piʝane sto 'meyaro] "Did they go to the Megara?"
7. [mazepse 'vroxonero] "Did it gather sewage?"
8. [ton 'gerðises sto 'domino] "Did you beat him at domino?"
9. [ferondan anayoyɔ] "Were they ill-mannered?"
10. [θa 'minune sta 'janena] "Will they live in Ioannina?"
11. ['ine apti 'mirina] "Is he from Mirina?"
12. [me'ʝalose sta 'meyara] "Did he grow up in Megara?"
13. [na 'piʝena 'monimu] "Should I have gone alone?"
14. [menune sto 'virona] "Do they live at Byron's?"
15. [ekane ka'rubalo] "Did he get a bump?"
16. [ɣno'rizis to xariðimo] "Do you know Charidimos?"

17. [ˈerixne ɔˈnonero] "Is it raining sleet?"
18. [se voiˈθa ɕe sto maˈyirema] "Does he also help cooking?"
19. [xriˈazete siðeˈroma] "Does it need ironing?"
20. [ˈine tu vlaðimiru] "Is it Vladimir's?"