Acoustic characteristics of Greek vowels produced by adults and children

Abstract

To our knowledge, no research on the acoustic analysis of Greek children's speech has been published. At the same time, the need for an expanded knowledge base, that would include information on children, is underlined in the literature (Deterding, 1990, Kent & Read, 1992, Clark & Yallop, 1995, Henton, 1995, Lass, 1996, Lee *et al*, 1997). This study, in an attempt to acknowledge this need, involves the acoustic analysis of Greek men's, women's and children's vowels, with an emphasis on the last, and examines the relationship between adults' and children's acoustic data. Children's vowels showed higher formant frequencies than adults', as expected from English children's data, and the scattergrams of adults' versus children's data revealed a relation that can be represented by a linear model quite satisfactorily.

Keywords: acoustics, formants, child speech, Greek vowels, vowel space, spectrography, regression

1. Introduction and previous studies

Acoustics is a recent field of study for the Greek language. The first acoustic studies of Greek vowels were published in the '80s, and since then there have been publications, but not many. Nevertheless, nothing has been published on the acoustics of Greek children's speech and this constituted the motivation for the present study.

The vowel system of Greek has been characterised a very common symmetrical one (Joseph & Philippaki-Warburton, 1987:263). A number of acoustic studies on Greek vowels have been carried out in the past. Kontosopoulos *et al.* (1988) analysed stressed and unstressed Greek vowels produced by seven male and seven female subjects. They commented that there is no overlap between articulatory spaces of neighbouring vowels and concluded that Greek vowels are distinct. Jongman *et al.* (1989) analysed stressed vowels produced by four male speakers and reported that Greek vowels are well separated in an acoustic space, allowing for maximal contrast between vowel categories.

This finding was also replicated in a perceptual study on the identification of synthetic stimuli by American and Greek subjects, carried out by Hawks *et al.* (1995). In this study, Greek listeners, unlike their American counterparts, rejected large numbers of stimuli as not possible Greek vowels. Another perceptual study on the identification of synthetic stimuli by four Greek female subjects was conducted by Botinis *et al.* (1997). There seemed to be very little or no overlap between categories on the perceptual vowel maps. If there was any overlap, it was between [a] and [o], while [I] and [u] were well separated for each subject.

The most recent acoustic study was published in 1999 by Fourakis *et al*. Their study investigated the effect of stress, tempo, focus and vowel quality on duration, F_0, amplitude and first and second formants of vowels produced by five Greek males. the results indicated that the Greek vowel system follows the universal tendencies in terms of duration but not in terms of F_0 and amplitude. Finally, Dauer (1981) found that the intrinsic factor of quality affects the duration of Greek vowels, hence high vowels are the shortest and low vowels are the longest. Also, vowels in stressed syllables are longer and have higher intensity. She also reported that the unstressed high vowels [i] and [u] are subject to extreme shortening, devoicing or elision in certain environments.

The present study looks at the first three formants of vowels produced by adults and children of both sexes, as well as the relationship between these values.

2. Methodology

2.1. The subjects

The subjects who took part in the experiment were 10 adult males, 10 females and 10 children of both genders. The adult male and female subjects were between 20 and 28 years old and were all undergraduate or postgraduate students at the University of Reading. Eight of the subjects came from Athens or central Greece, nine of the subjects came from Northern Greece (mainly Thessaloniki) and 3 subjects came from Southern Greece (Crete). All subjects spoke standard Greek[1]. All the children came from Crete and their ages ranged from five to ten. Three of them were male and seven female. They all were monolingual speakers of Greek with no reported history of speech, language or hearing problems.

[1] Although each area has certain "accent characteristics", these are not likely to be detected in an acoustic analysis of vowels produced by reading from a list. Hence, these differences do not constitute a limitation to this acoustic study of vowels. It might be interesting, though, to confirm this opinion. According to

2.2. The corpus

The corpus consisted of 30 words. The set of the first six words contained the vowel [i] in three positions: word-initially (e.g. ìrθa, imèra) word-medially (e.g. poðilato, iðikòs) and word-finally (e.g. cerì, mesimèri). The first word of each pair contained the stressed vowel, whereas the second word of the pair contained the unstressed vowel. The set of the next six words contained the vowel [e] in the same positions and stress conditions. As all words had more than one syllable (up to four syllables), they contained more than one vowel in several positions, stressed or unstressed. These instances were also taken into account, so that the analysis would be based on more tokens of vowels, rather than only six for each vowel. So, the final number of tokens for each vowel was: 15 for [i], 17 for [e], 23 for [a], 19 for [o] and 6 for [u]. In total, there were 80 tokens of vowels in the aforementioned positions and stress conditions. It is noteworthy that [u] did not occur in any of the other words, apart from the ones especially selected to contain it (last set of six words). This demonstrates the very low frequency of occurrence of this vowel in Greek (Dauer, 1981:20).

2.3. The recording technique

Two different recording techniques were used. The recordings of the adult subjects were made on a portable minidisc recorder and a hyper-cardioid dynamic

---

Mackridge (1985:5-6), “the speech (even the pronunciation) of moderately educated people from all parts of Greece tends to be hardly distinguishable from that of an Athenian”.

microphone was held within 30 cm from their mouth. The recording level was set separately for each individual and was kept constant throughout the recording. The recordings of the children were made on a stereo recorder and saved on an audio-cassette. A lapel microphone was used, so as to keep the distance fixed and avoid overloading. The children who felt confident with their reading, read the list, whereas those who could not read, were asked to repeat each word after their parent. Two of the children, aged five and six, went for the second option. Each subject was recorded twice and the best recording was chosen for analysis.

2.4 Measurement technique and methods of analysis

The recordings were transferred to the Kay Elemetrics Computerized Speech Lab (CSL) and the sampling rate was set at 8 kHz. The first three formants were measured from wide-band spectrograms. Since there were 80 vowel tokens for each subject and 30 subjects were recorded, 2,400 vowels were analysed on total. The formants were extracted at approximately the temporal midpoint of the vowel to avoid influences from the surrounding consonants. An FFT-derived spectrum taken at the middle of the vowel supplemented the spectrogram measurement in cases where formants could not be determined with certainty. If a frequency value still could not be estimated with certainty, it was left out. This resulted in the appearance of missing values in the data. the highest number of missing values concerns the third formant (F_3) of [o] and [u] produced by children. This might have to do with the fact that the fundamental frequency of children's voices is high, thus creating problems in formant location. F_3 was also often low in intensity in children's speech. I.e. in [u], F_3 was absent in non-stressed, word-final position; F_1 and F_2 were also hard to discern, and, in some cases, there was no trace of the vowel on the waveform.

For the analysis of the data the MINITAB statistical software program was used. Descriptive Statistics was performed for each one of the three formants of the five vowels produced by the three different classes of speakers. The effect of stress and position in word was not investigated at this point, as we were interested in getting a more general picture of the vowels' formants in all environments. Regression analysis was also used to find the regression equation which would predict child formant values from adult ones.

2. Results and discussion

Table 1 shows the mean values of the first three formants of the five Greek vowels produced by men, women and children.

2.1. Comparisons between classes of speakers

The results are presented comparatively for the three classes of speakers (adult male, adult female and child); first in figures 1-5, for each vowel separately, and then in

i			
	Male	**Female**	**Child**
F_1	423	469	581
$\mathbf{F_2}$	2073	2571	2873
$\mathbf{F_3}$	2593	3109	3637
e			
	Male	**Female**	**Child**
F_1	601	687	719
$\mathbf{F_2}$	1811	2231	2607
$\mathbf{F_3}$	2560	3051	3612
a			
	Male	**Female**	**Child**
F_1	736	873	922
$\mathbf{F_2}$	1466	1699	1811
$\mathbf{F_3}$	2459	2713	3197
o			
	Male	**Female**	**Child**
F_1	583	657	730
$\mathbf{F_2}$	1137	1219	1462
$\mathbf{F_3}$	2479	2817	3240
u			
	Male	**Female**	**Child**
F_1	434	451	560
$\mathbf{F_2}$	921	955	1190
$\mathbf{F_3}$	2460	2804	3172

Table1 The mean values of F_1, F_2 and F_3 (in Hz) of the five Greek vowels produced by men, women and children.

fig.6, for all vowels. In figures 1-5, F1 is plotted against F2, so that the distribution values of the three classes of speakers can be compared.

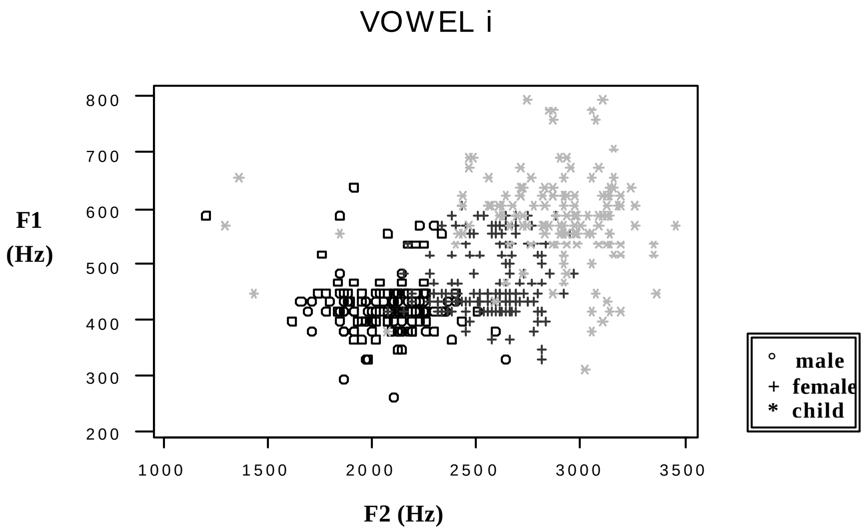

Fig. 1 Plot of vowel [i], showing the distribution of the male, female and child formant frequency data.

Fig. 1 shows the distribution values for vowel [i]. The emerging pattern is that adult male formants are the lowest in frequency, adult female formants are intermediate and child formants are the highest. We can discern only very few child formants in the adult male 'region'. The same pattern arises for vowel [e], as we can see in fig. 2.

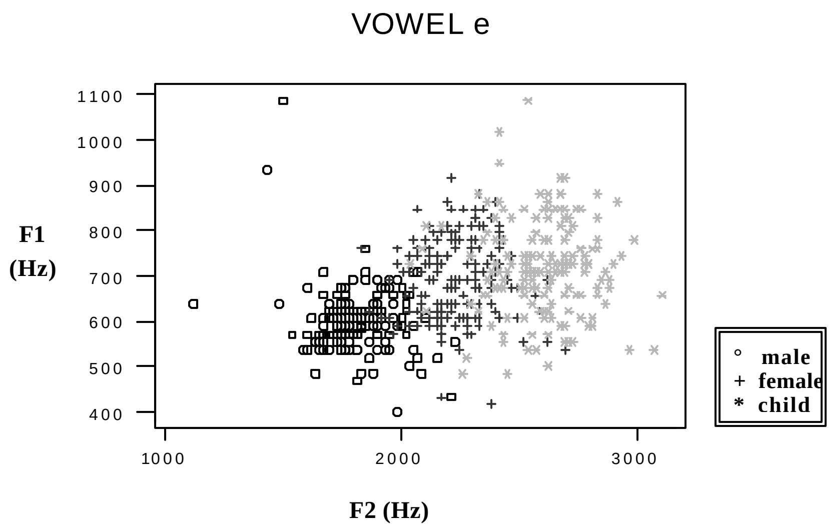

Fig.2 Plot of vowel [e], showing the distribution of the male, female and child formant frequency data.

The picture changes for vowel [a] in fig.3. Although adult male and female values maintain the same pattern, the child values seem more scattered, so that many child values a are amongst adult values.

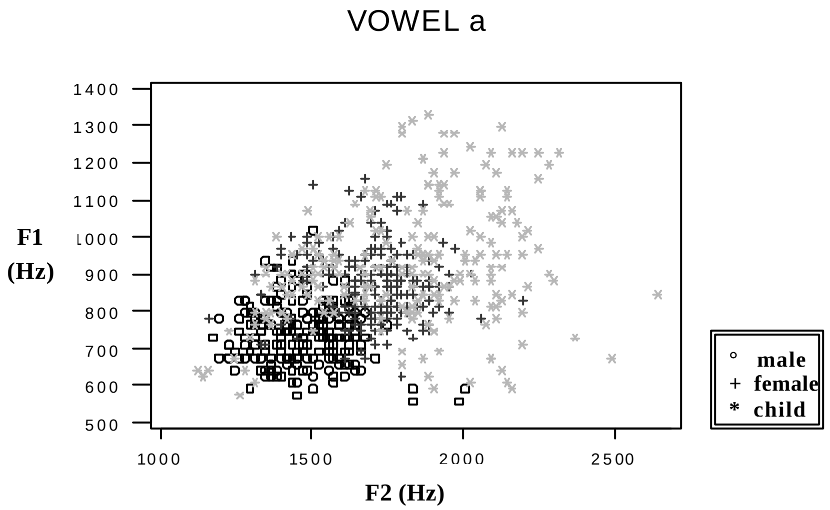

Fig.3 Plot of vowel [a], showing the distribution of the male, female and child formant frequency data.

The same observation can be made for vowel [o] in fig.4. F2 for all classes is lower but some child values are quite high. The problem concerning missing values due to formant location difficulties is evident in vowel [u], fig.5. Although the values seem scattered, the basic pattern is maintained and child formants are generally higher.

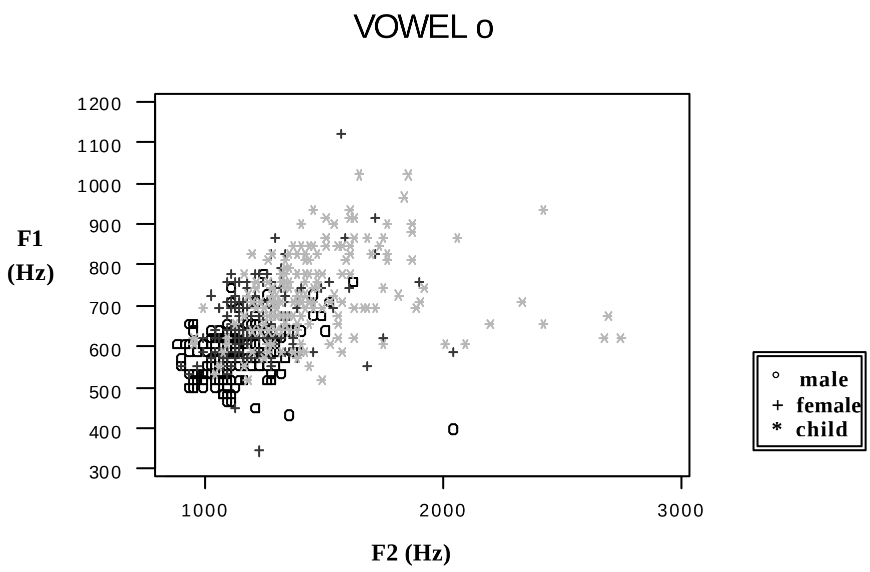

Fig.4 Plot of vowel [o], showing the distribution of the male, female and child formant frequency data.

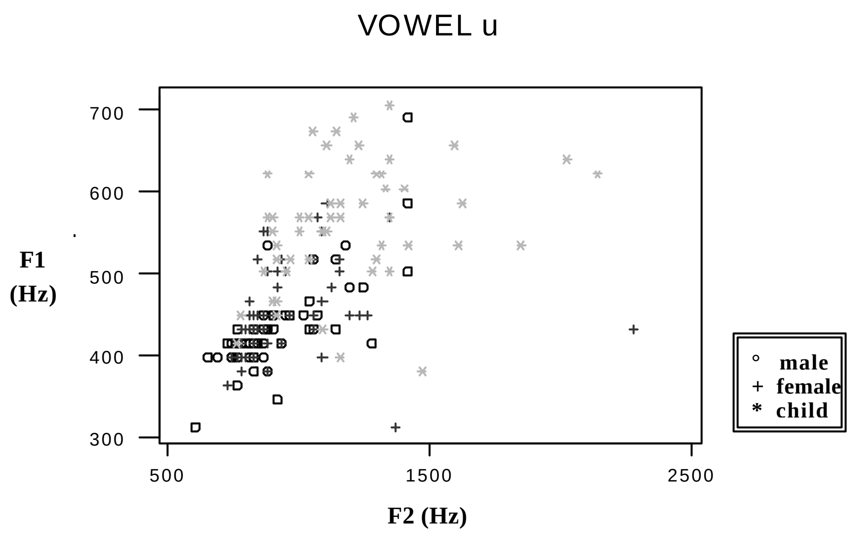

Fig.5 Plot of vowel [u], showing the distribution of the male, female and child formant frequency data.

Concerning the general picture of all vowels in fig.6, child F1 and F2 tend to be higher than adult formants. Hence, the child vowel space is placed more 'downwards and to the left' in comparison with the adult vowel spaces. This vowel space pattern for the three classes of Greek speakers has many similarities with Deterding's findings concerning English speakers' vowel spaces (Deterding, 1990:49-51). It is also noteworthy that the women's vowel space is the broadest, child space follows and the men's is the smallest, which is also the case with English (Deterding, 1990:47-49). We also observe that the F2 difference between child and adult values is higher for front vowels [i, e] than for back vowels [o, u].

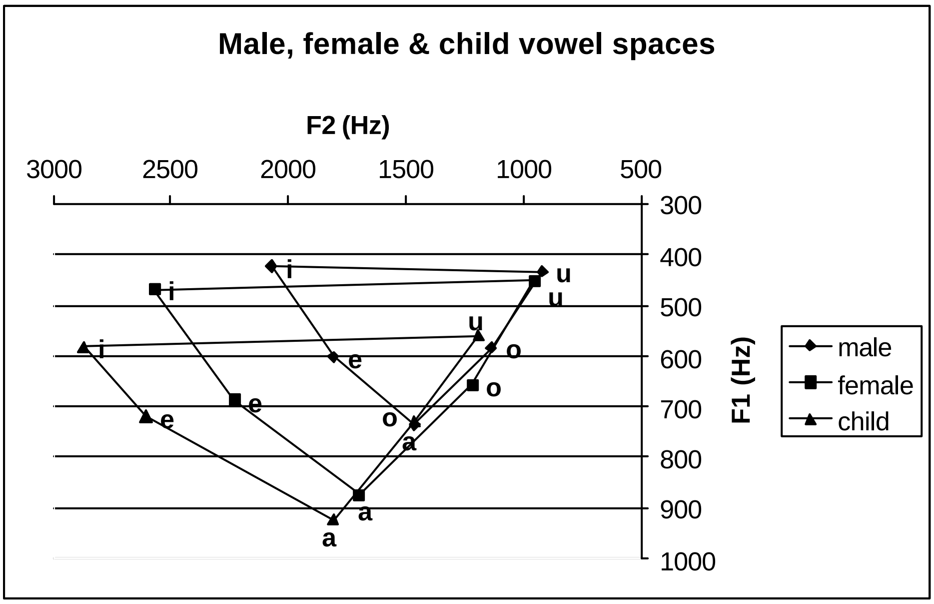

Fig.6 Greek adult male, female and child vowel spaces

2.2. Relationships between adult male, adult female and children formants

An interesting theoretical question is the nature of the relationship between adult male and female formant frequencies and between adult and child formant frequencies. This issue has also a practical side, as the investigation of such relationships would indicate a way in which differences could be normalised, thus constituting the base for a methodology to achieve more accurate speech recognition systems. The problem of automatic recognition of children’s speech has gained attention in the recent years and it has been shown that speech-recognition systems trained on adult speech perform unsatisfactorily when tested on children’s speech (Lee *et al*, 1999:1455).

In an attempt to investigate the above relationships, we plotted male versus female formant frequencies in fig.7(a). We observe that the data tend to create a line. The same observation is made when we plot male versus child data (fig.7(b)) and female versus child data (fig.7(c)). So there is some degree of linear relation between the variables and hence this relation can be represented by a linear model.

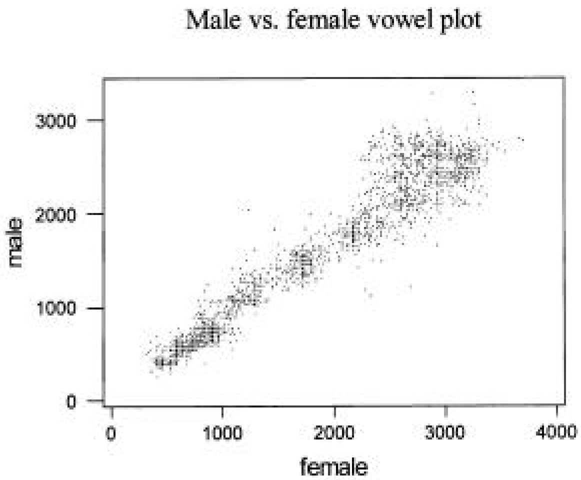

Fig.7(a) Adult male vs. adult female data

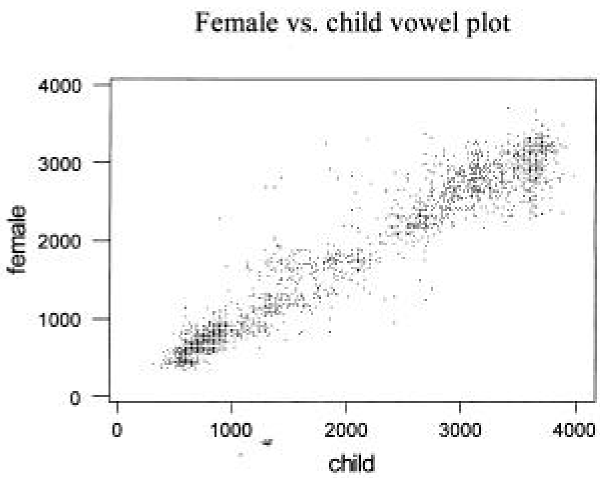

Fig.7(b) Adult female vs. child data

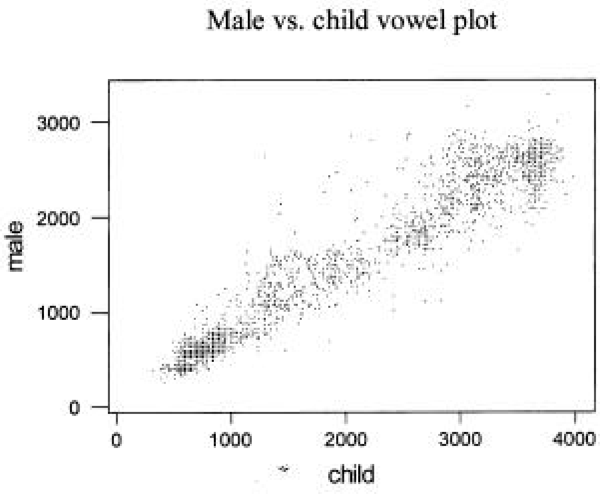

Fig.7(c) Adult male vs. child data

We used multiple regression analysis to find an equation which will help to explain the variation in one variable (child formants) by using two other variables (male and female adult formants). Table2 shoes the regression equations predicting child vowel formants from adult values. Because the R-Sq value is higher than 90% in all cases, it is obvious that a large amount of this variation is explained by this equation.

Vowel	Equation	R-Sq
[i]	child[i] = 43.6 + 0.482 male[i]+ 0.731 female[i]	96.3%
[e]	child[e] = - 88.7 + 0.369 male[e] + 0.891 female[e]	97.7%
[a]	child[a] = - 93.4 + 0.760 male[a] + 0.506 female[a]	90.5%
[o]	child[o] = 13.8 + 0.631 male[o] + 0.579 female[o]	93.8%
[u]	child[u] = 42.6 + 0.963 male[u] + 0.262 female[u]	91.1%
All	**child[v] = - 39.4 + 0.568 male[v] + 0.678 female[v]**	94.7%

Table2 Equations predicting child formant values using adult values

3. Conclusions

This study investigates the first three formants of vowels produced by adults and children of both sexes, as well as the relationship between these values.

- Many difficulties are encountered when trying to locate F_3 frequencies of children's back vowels.
- Child formants are found to be higher than adults', so that the child vowel space is placed more 'downwards and to the left' in comparison with the adult vowel spaces.
- The speaker class pattern of vowel space for Greek is similar to the English one (Deterding, 1990).
- The child formant values can be predicted quite satisfactorily (R-Sq>90%) using adult formant values.

References

Botinis, A., Fourakis, M. and J. W. Hawks (1997). "A perceptual study of the Greek vowel space using synthetic stimuli". In *Eurospeech '97: Proceedings of the 5th European Conference on Speech Communication and Technology*, Patras: University of Patras.

Botinis, A., Fourakis, M. and M. Katsaiti (1995). "Acoustic characteristics of Greek vowels under different prosodic conditions". In *Proceedings of the XIIIth International Congress of Phonetic Sciences* (vol. 4). 404-407.

Clark, J. and C. Yallop (1995). *An introduction to phonetics and phonology*. 2nd edition. Oxford: Blackwell.

Dauer, R. (1981). “The reduction of unstressed high vowels in modern Greek”. *Journal of the International Phonetic Association* 10: 17-27.

Deterding, D. H. (1990). *Speaker normalisation for automatic speech recognition.* Ph.D. Thesis. University of Cambridge.

Fourakis, M., Botinis, A. and M. Katsaiti (1999). “Acoustic characteristics of Greek vowels”. *Phonetica* 56: 28-43.

Hawks, J.W. and M.S. Fourakis (1995). “The perceptual vowel spaces of American English and modern Greek: a comparison”. *Language and Speech* 38: 237-252.

Henton, C. (1995). “Cross-language variation in the vowels of female and male speakers”. In *Proceedings of the XIIIth International Congress of Phonetic Sciences* (vol. 4). 420-423.

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

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

Kent, R.D. and C. Read (1992). *The acoustic analysis of speech*. San Diego: Singular Publishing Group.

Kontosopoulos, N., Ksiromeritis, N. and A. Tsitsa (1988). “Akustikes erevnes epi ton elinikon fthogon (Acoustic research on the Greek sounds)”. *Glossologia* 5-6: 103-122.

Lass, N. J. (ed.) (1996). *Principles of experimental phonetics*. London: Mosby.

Lee, S., Potamianos, A. and S. Narayanan (1999). “Acoustics of children’s speech: Developmental changes of temporal and spectral parameters”. *Acoustical Society of America* 105: 1455-1468.

Mackridge, P. (1985). *The modern Greek language. A descriptive analysis of standard modern Greek.* Oxford: OUP.