

## Acoustic characteristics of Greek vowels produced by adults and children

*Anna Sfakianaki*

### Abstract

To my 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, which would include information on children, is underlined in the literature (Deterding 1990, Kent and Read 1992, Clark and Yallop 1995, Henton 1995, Lass 1996, Lee *et al.* 1997). In response to this need, this study 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 are found to show higher formant frequencies than adults', as expected from English children's data, and the scattergrams of adults-versus-children's data reveal a relation that can be represented by a linear model quite satisfactorily.

**Key words:** acoustics, formants, child speech, Greek vowels, vowel space, spectrography, regression

### 1. Introduction

Acoustics is a recent field of study for the Greek language. The first relevant studies of Greek vowels were published in the '80s, and since then there have been some 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 as a very common symmetrical one (Joseph and Philippaki-Warburton 1987: 263). A number of

acoustic studies on Greek vowels have been carried out in the past. For instance, Κοντοσόπουλος, Ξηρομερίτης and Τσίτσα (1988) analysed stressed and unstressed Greek vowels, produced by seven male and seven female subjects. They found out that there is no overlap between articulatory spaces of neighbouring vowels and concluded that Greek vowels are distinct. Jongman, Fourakis and Sereno (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 and Fourakis (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, Fourakis and Hawks (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, Botinis and Katsaiti. 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 ten (10) adult males, ten (10) females and ten (10) children of both sexes. The adult male and female subjects were between 20 and 28 years old, and were all undergraduate or post-

graduate students at the University of Reading. Eight (8) of the subjects came from Athens or central Greece, nine (9) of the subjects came from Northern Greece (mainly Thessaloniki), and three (3) subjects came from Southern Greece (Crete). All subjects spoke standard Greek.<sup>1</sup> 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.

## 2.2 *The corpus*

The corpus consisted of thirty (30) words. The set of the first six (6) words contained the vowel [i] in three positions: word-initially (e.g. 'irθa, i'mera) word-medially (e.g. po'ðilato, iðikos) and word-finally (e.g. ce'ri, mesi'meri). The first word of each pair contained the stressed vowel, whereas the second one contained the unstressed vowel. The set of the next six (6) words contained the vowel [e] in the same positions and stress conditions, the next six (6) the vowel [a], and so forth.

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 on 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 eighty (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 (6) 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 microphone was held within 30cm from their mouth. The recording level was set

---

<sup>1</sup> 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 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".



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 audiocassette. 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 in 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 which 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. For instance, 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 I was 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.

### **3. 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.

**Table 1.** The mean values of  $F_1$ ,  $F_2$  and  $F_3$  (in Hz) of the five Greek vowels produced by men, woman and children.

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

### 3.1 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 Figure 6, for all vowels. In Figures 1-5,  $F_1$  is plotted against  $F_2$ , so that the distribution values of the three classes of speakers can be compared.

All children, regardless of sex and age, have been grouped together as one speaker class. Although looking into the development of acoustic parameters in Greek children's speech according to age and sex is a very interesting subject, it

is outside the scope of the present study. This study attempts to demonstrate fundamental differences and relationships between the above three classes of speakers in terms of vowel formants and acoustic vowel space. Furthermore, longitudinal studies and research based on large speech databases of children with varying age of both sexes in American English have shown that differentiation of male and female formant patterns begin around the age ten or eleven, whereas the subjects of this study are up to ten years old, and that children's speech acoustic characteristics fully develop to adult level around age fifteen (Lee *et al.* 1999).

In addition, creating different subgroups according to children's age would not prove very helpful. Smith and Kenney (1998) found that various acoustic parameters of children's speech do not necessarily develop in comparable ways or at similar rates across children and within the same child. This makes a categorisation according to age no more useful than looking at the whole picture regardless of age, as acoustic characteristics are expected to show considerable variation across individuals anyway.

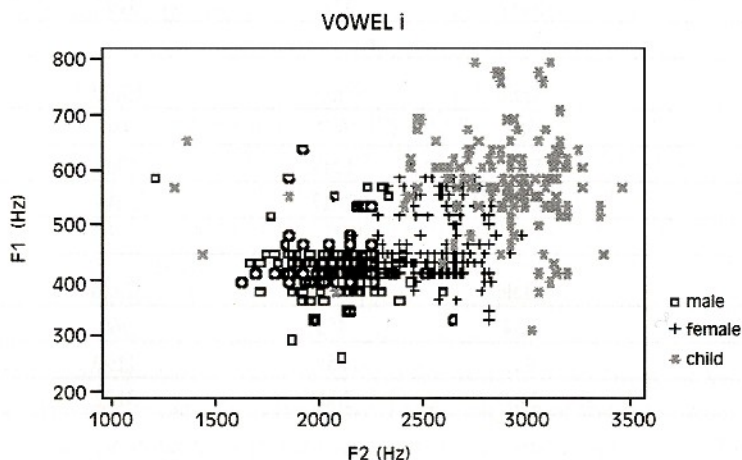
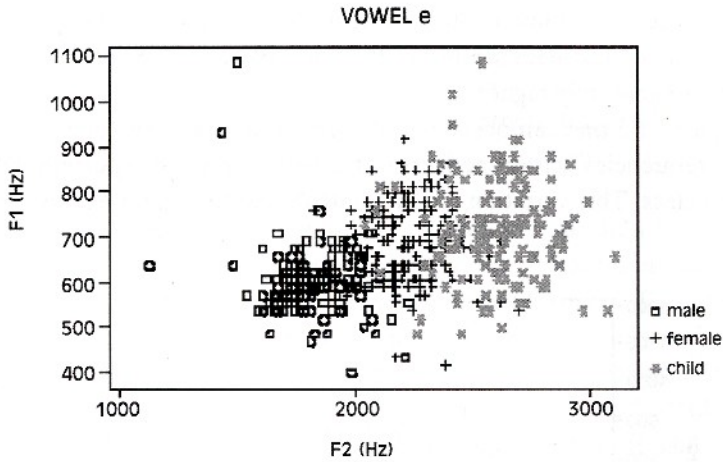


Figure 1. Plot of vowel [i], showing the distribution of the male, female and child formant frequency data.

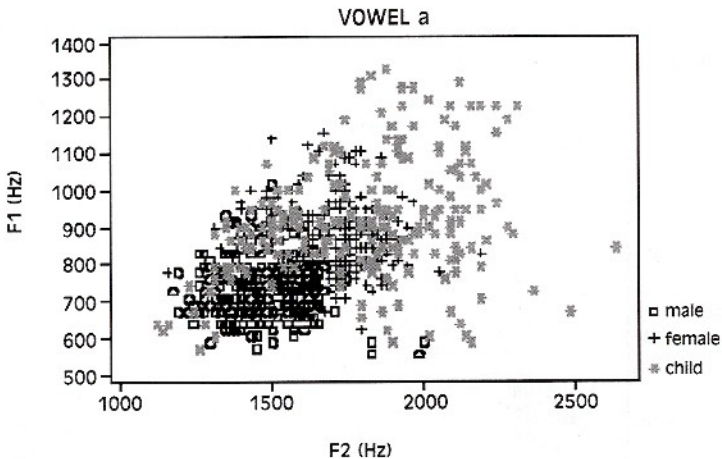
Figure 1 shows the distribution values for vowel [i]. The emerging pattern is that adult male formants are the lowest in frequency, adult female formants are intermediate and child formants are the highest. We can discern only very few child formants in the adult male 'region'. The same pattern arises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





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

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



*Figure 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 Figure 4.  $F_2$  for all classes is lower but some child values are quite high. The problem concerning miss-

ing values, due to formant location difficulties is evident in vowel [u], Figure 5. Although the values seem scattered, the basic pattern is maintained and child formants are generally higher.

In Figures 1-5 one can observe that there is variation between the values of formant frequencies across speakers of the same class and especially within the children's class. This variation could be attributed to a great extent to the fact

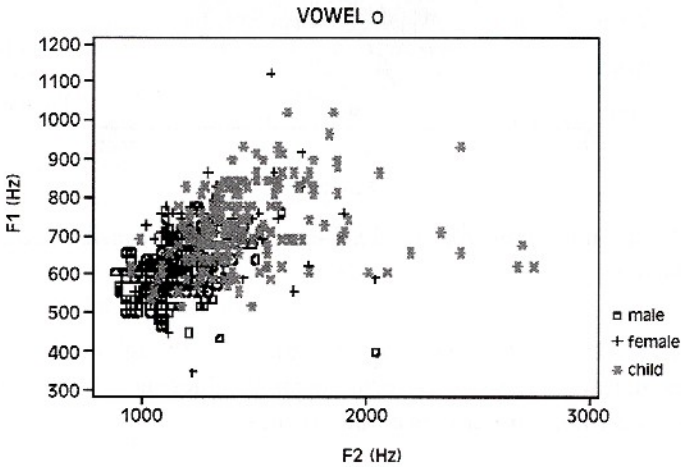


Figure 4. Plot of vowel [o], showing the distribution of the male, female and child formant frequency data.

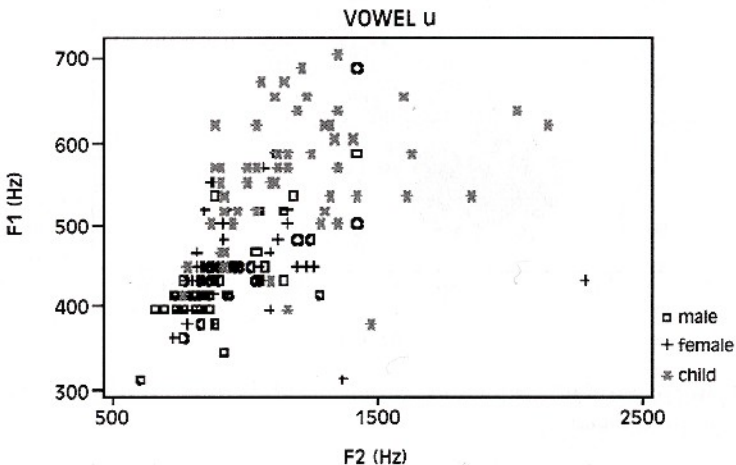


Figure 5. Plot of vowel [u], showing the distribution of the male, female and child formant frequency data.



that the rate of development is different for different children (Smith and Kenney 1998). In addition, Lee, Potamianos and Narayanan (1999) found out that children younger than ten demonstrate temporal and spectral variability which suggests that they have not fully established their own optimal articulatory vowel targets in a given context. Thus, children's formants are bound to create a more scattered picture as in Figures 1-5.

As regards the general picture of all vowels in Figure 6, child  $F_1$  and  $F_2$  tend to be higher than adult formants. Hence, the child vowel space is displaced '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 (: 47-49). We also observe that the  $F_2$  difference between child and adult values is higher for front vowels [i, e] than for back vowels [o, u].

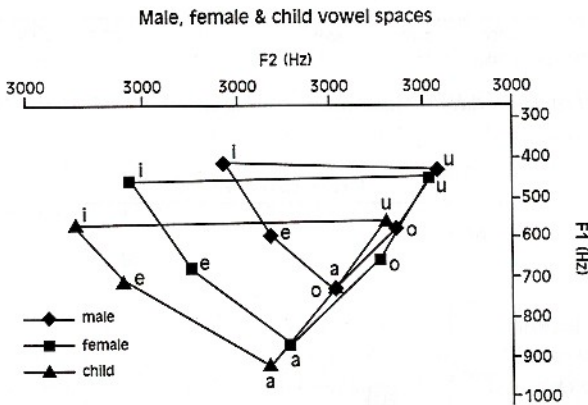


Figure 6. Greek adult male, female and child vowel spaces.

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

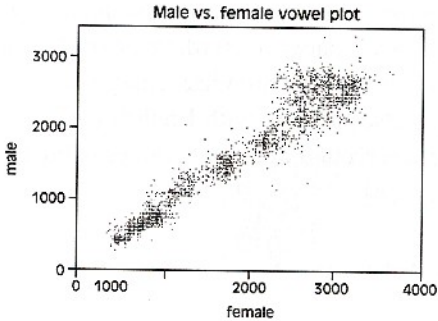


Figure 7a. Adult male vs. adult female data

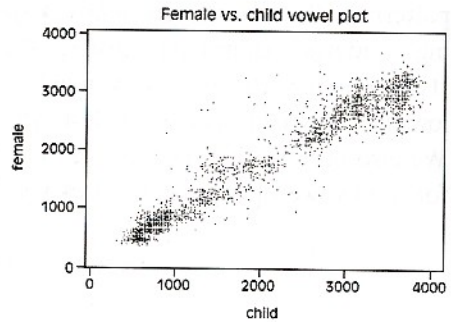


Figure 7b. Adult female vs. child data

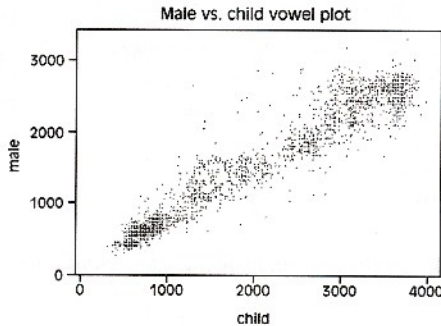


Figure 7c. Adult male vs. child data

An equation was calculated by a multiple regression analysis, which would help to explain the variation in one variable (child formants), by using two other variables (male and female adult formants). Table 2 shows 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.

**Table 2.** Equations predicting child formant values using adult values.

VOWEL	EQUATION	R-Sq
[i]	$\text{child}[i] = 43.6 + 0.482 \text{ male}[i] + 0.731 \text{ female}[i]$	96.3%
[e]	$\text{child}[e] = -88.7 + 0.369 \text{ male}[e] + 0.891 \text{ female}[e]$	97.7%
[a]	$\text{child}[a] = -93.4 + 0.760 \text{ male}[a] + 0.506 \text{ female}[a]$	90.5%
[o]	$\text{child}[o] = 13.8 + 0.631 \text{ male}[o] + 0.579 \text{ female}[o]$	93.8%
[u]	$\text{child}[u] = 42.6 + 0.963 \text{ male}[u] + 0.262 \text{ female}[u]$	91.1%
ALL	$\text{child}[v] = -39.4 + 0.568 \text{ male}[v] + 0.678 \text{ female}[v]$	94.7%

#### 4. Conclusions

This study has investigated the first three formants of vowels produced by adults and children of both sexes, as well as the relationship between these values. On the whole, the speaker class pattern of vowel space for Greek has been shown to be similar to the English one (Deterding 1990). Adult male vowels have been found to be the lowest, while child formants the highest, and so the child vowel space is displaced 'downwards and to the left', in comparison with the adult vowel space.  $F_3$  frequencies of children's back vowels have proved difficult to locate. Finally, child formant values can be predicted quite satisfactorily ( $R\text{-Sq} > 90\%$ ), using adult formant values in a multiple regression analysis.

#### References

- Botinis, Antonis, Marios Fourakis and John 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. (on CD ROM).
- Clark, John and Colin Yallop (1995). *An Introduction to Phonetics and Phonology*. Oxford: Blackwell.
- Dauer, Rebecca (1981). "The reduction of unstressed high vowels in Modern Greek". *Journal of the International Phonetic Association* 10: 17-27.
- Deterding, David H. (1990). *Speaker Normalisation for Automatic Speech Recognition*. Ph.D. Thesis. University of Cambridge.
- Fourakis, Marios, Antonis Botinis and Maria Katsaiti (1999). "Acoustic characteristics of Greek vowels". *Phonetica* 56: 28-43.
- Hawks, John W. and Marios S. Fourakis (1995). "The perceptual vowel spaces of American English and Modern Greek: a comparison". *Language and Speech* 38: 237-252.
- Henton, Caroline (1995). "Cross-language variation in the vowels of female and male speakers". In *Proceedings of the XIII<sup>th</sup> International Congress of Phonetic Sciences* Vol. 4, 420-



423. Stockholm.

- Jongman, Allard, Marios Fourakis and Joan Sereno (1989). "The acoustic vowel space of Modern Greek and German". *Language and Speech* 32: 221-248.
- Joseph, Brian D. and Irene Philippaki-Warbuton (1987). *Modern Greek*. London: Croom Helm.
- Kent, Ray D. and Charles Read (1992). *The Acoustic Analysis of Speech*. San Diego: Singular Publishing Group.
- Lass, Norman J. (ed.) (1996). *Principles of Experimental Phonetics*. London: Mosby.
- Lee, Sungbok, Alexandros Potamianos and Shrikanth Narayanan (1999). "Acoustics of children's speech: Developmental changes of temporal and spectral parameters". *Acoustical Society of America* 105: 1455-1468.
- Mackridge, Peter (1985). *The Modern Greek Language. A Descriptive Analysis of Standard Modern Greek*. Oxford: Oxford University Press.
- Smith, Bruce L. and Mary K. Kenney (1998). "An assessment of several acoustic parameters in children's speech production development: Longitudinal data". *Journal of Phonetics* 26: 95-108.
- Κοντοσόπουλος Νικόλαος, Νικόλαος Ξηρομερίτης και Αναστασία Τσίτσα (1988). «Ακουστικές έρευνες επί των ελληνικών φθόγγων» [Acoustic research on the Greek sounds]. *Γλωσσολογία* 5-6: 103-122.