A PERCEPTUAL STUDY OF THE GREEK VOWEL SPACE USING SYNTHETIC STIMULI

A. Botinis, M. Fourakis**, and J. W. Hawks****
*Linguistics Department, Athens University
** Department of Speech and Hearing Science, The Ohio State University
*** School of Speech Pathology and Audiology, Kent State University
E-mail: fourakis.1@osu.edu

ABSTRACT

Four female native speakers of Modern Greek listened to 465 synthetic vowel tokens with F1 frequencies ranging from 250 to 800 Hz and F2 frequencies ranging from 900 to 2900 Hz in 50 Hz steps. They were asked to identify each stimulus as one of the five vowels of Modern Greek or to reject it if they thought it could not be a vowel of their language. The subjects rejected about 64 percent of the tokens as not possible vowels. The remaining points were plotted in an F1 by F2 space with the codes assigned by each subject and in a composite space, where only the points identified with the same response by at least three subjects were used. The results replicated those of Hawks and Fourakis [1], except that the code for the vowel [e] was assigned to many more points than the codes for the other vowels.

1. INTRODUCTION

Hawks and Fourakis [1] reported on the identification of synthetic vowel stimuli by American and Greek subjects. The subjects listened to 304 synthetic vowel stimuli and were asked to either reject a stimulus as not a possible vowel in their language or identify it as one of the vowel categories of their language. They found that the Greek listeners, unlike their American counterparts, rejected large numbers of stimuli as not possible vowels for Greek. The stimuli which were identified as possible Greek vowels, with at least five of the ten subjects agreeing on the vowel category that was assigned to the stimulus, formed well-defined and separated subspaces in the vowel space. They interpreted these results as implying that the Greek vowel were organized in a maximally contrasting manner as proposed by Liljencrants and Lindblom [2]. In this current work we examine the responses of Greek subjects to synthetic vowel stimuli that span the vowel space with F1 frequencies ranging from 250 to 800 Hz and F2 frequencies ranging from 900 to 2900 Hz. The total number of stimuli is thus 465, more than 50% larger than the number used in Hawks and Fourakis [1], providing for much finer resolution of the vowel space. In addition, the data reported here are from female subjects only. In the full version of this paper, data will be reported from male and child subjects.

2. METHOD

2.1. Stimuli

Synthetic vowel tokens were constructed using center formant frequencies ranging from 250 to 800 Hz for F1 and from 900 to 3000 Hz for F2, in 50 Hz steps for both formants. The resulting quadrangle would thus include 516 different vowel tokens. The frequency of F3 was computed on the basis of F1 and F2 frequencies using the algorithm described in Nearey [3] This algorithm sometimes computed F3 frequencies higher than the frequency set for F4 (3700 Hz). This occurred in 51 tokens which were eliminated from further consideration, reducing the upper F2 frequency to 2900 Hz and the number of tokens used to 465. Bandwidths for each formant were

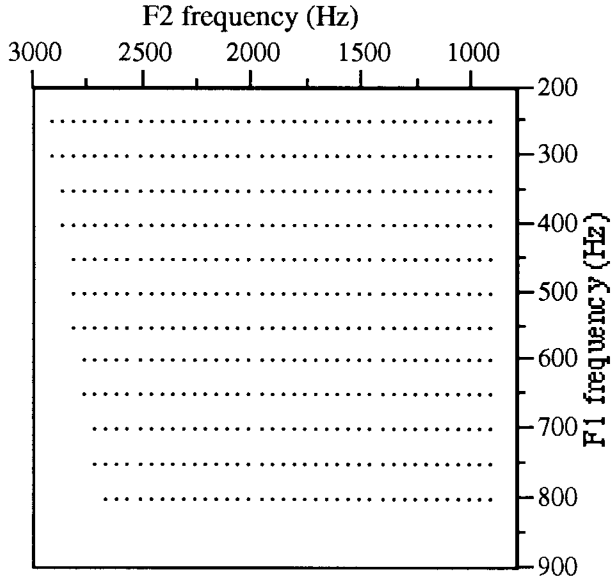

Fig 1. All F1 and F2 frequency pairs synthesized and used in this experiment.

computed using the algorithm described in Hawks and Miller [4]. The synthesis employed the cascade branch of the KLSYN88A software synthesizer, Klatt and Klatt [5], using a10 KHz sampling rate, with 12 bit precision. All tokens were 250 ms long, with 20 ms amplitude on and off ramps. Overall amplitude was normalized across tokens within +/- 1 dB. The frequency of F4 was fixed at 3700 Hz. Figure 1 shows all the F1 - F2 pairs synthesized for this experiment. The stimuli were converted to the file format used by the Computerized Speech Lab (CSL) and played out through a speaker at a comfortable level. CSL provides its own anti-aliasing filter. A different randomized order was used for each subject.

2.2. Subjects and data collection

The subjects were four female, undergraduate students at the University of Athens Linguistics department. They were asked to listen to each vowel token and respond in one of two ways: If they thought the token could be considered a vowel of Modern Greek, they indicated the vowel on their response sheet. If not, they underlined the word "oxi" (no). For each token and decision, they were also asked to indicate their certainty by assigning a certainty rating ranging from 5 (very certain) to 1 (least certain).

3. RESULTS

3.1. Rejection rates

The rejection rate, that is the percentage of tokens judged as not possible vowels of Modern Greek, ranged from 55% for subject FS3 to 81 percent for subject FS2. Across subjects the rejection rate was 64%, This figure is very close to that reported in Hawks and Fourakis [1], which was about 65% for ten subjects.

3.2. Certainty ratings

Subjects varied in how certain they were about their responses. The average certainty response for subject FS1 was 4.48 while for FS2 the average was 2.57. Subjects FS3 and FS4 had averages of 3.72 and 3.49 respectively. The perceptual vowel maps that were created used tokens with ratings higher than three for all subjects except FS2, who assigned very few certainty ratings above 3, and most of those were for the "not vowel of Greek" response.

3.3. Perceptual vowel maps

Figures 2-5 show perceptual vowel maps constructed for each subject. Each F1-F2 frequency pair identified as a possible vowel of Modern Greek with a certainty rating of 3 or more (except for subject FS2) is plotted by the category symbol assigned by the subject. Overall it can be seen that points assigned to the [e] category are more numerous than any of the other categories, and this is especially true for subjects FS1 and FS3. This category thus occupies much more space than the others. For each subject there seems to be little or no overlap between categories. Most of the overlap, when

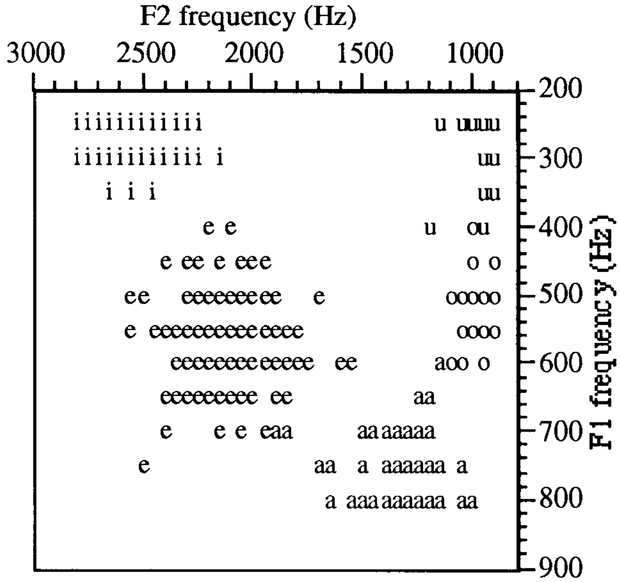

Fig. 2. Subject FS1 responses with certainty higher than 3.

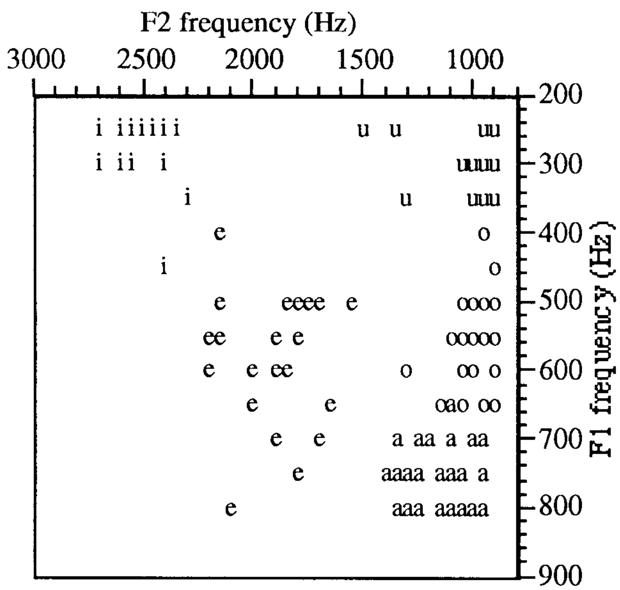

Fig. 3. All responses for Subject FS2.

there is any, is between [a] and [o], while [i] and [u] are well separated for each subject.

Figure 6, shows the perceptual vowel map constructed out of those points for which at least three subject assigned the same category code. This composite map shows that all the categories are well separated and that large portions of the space are unoccupied, reflecting the high average

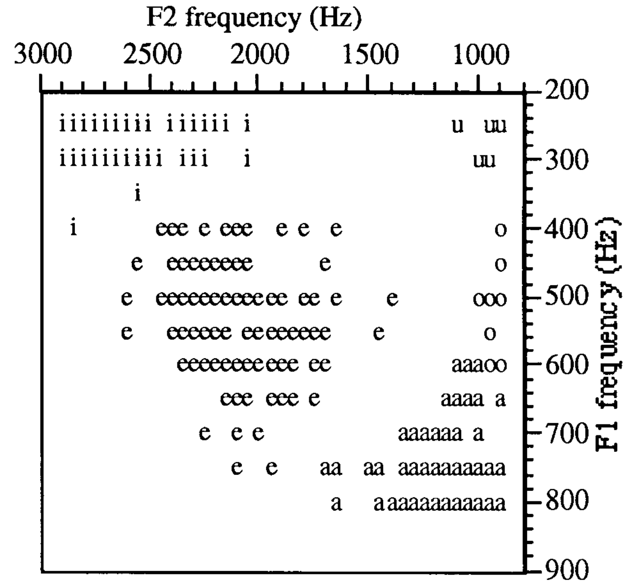

Fig. 4. Subject FS3 responses with certainty higher than 3.

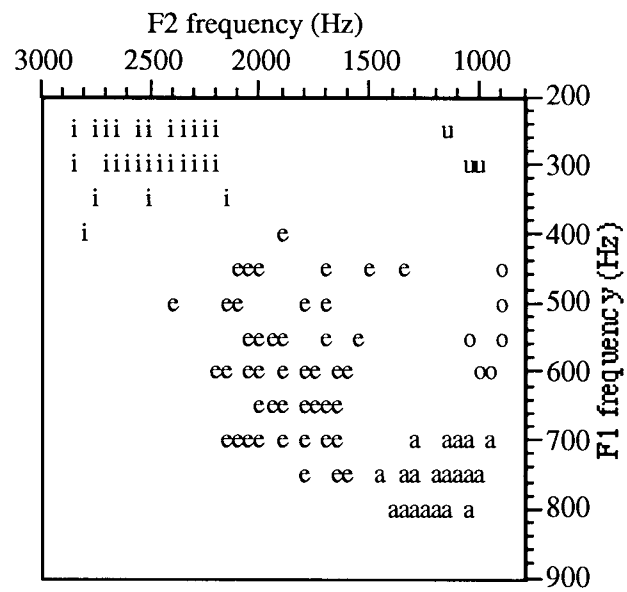

Fig. 5. Subject FS4 responses with certainty higher than 3.

rejection rate. This map is very similar to the one reported by Hawks and Fourakis [1].

4. DISCUSSION

The results of the present experiment replicated Hawks and Fourakis [1]. The rejection rate was very high (64%) and the vowel spaces

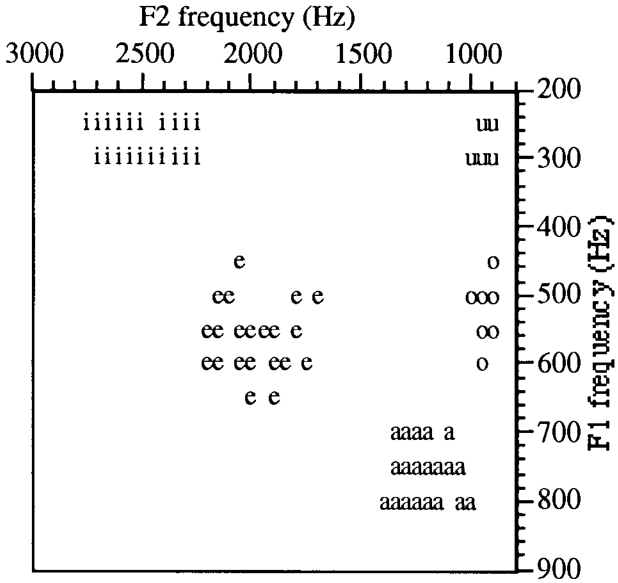

Fig. 6. All points for which 3 out of 4 subjects agreed on the identification code assigned.

constructed for each subject, as well as the composite vowel space, showed well defined categories with very little overlap. The main difference between these results and those of Hawks and Fourakis lies in the number of tokens assigned to the [e] category, which was much larger than previously reported. Overall, the results support the claim that the Greek vowel space, at least as it can be constructed form subjects' responses to synthetic stimuli, is organized in a maximally contrastive manner, Liljencrants and Lindblom [2].

5. REFERENCES

[1] J. W. Hawks and M. Fourakis, "The Perceptual Vowel Spaces of American English and Modern Greek: A Comparison", *Language and Speech*, Vol. 38, pp. 237-252, 1995.

[2] J. Liljencrants and B. Lindblom, "Numerical Simulation of Vowel Quality Systems: The Role of Perceptual Contrast', *Language*, Vol. 48, pp. 839-862, 1972.

[3] T. Nearey, "Static, Dynamic, and Relational Properties in Vowel Perception", *J. Acoust. Soc. Am.*, Vol. 85, pp. 2088-2113, 1989.

[4] J. W. Hawks and J. D. Miller, "A Formant Bandwidth Estimation Procedure for Vowel Synthesis", *J. Acoust. Soc. Am.*, Vol. 97, pp. 1343-1344, 1995

[5] D. Klatt and L. C. Klatt, "Analysis, Synthesis, and Perception of Voice Quality Variations among Female and Male Talkers", *J. Acoust. Soc. Am.*, Vol. 87, pp. 820-857, 1990.

Acknowledgments

The authors would like to thank Irene Prinou, Department of Linguistics, Athens University, for help with running the subjects and collecting the data.