Perception of vowels across Greek dialects

ANGELOS LENGERIS
University of Kent

1. Introduction

According to the Theory of Adaptive Dispersion, vowels are maximally dispersed and evenly spaced so as to enhance contrasts (Liljencrants & Lindblom 1972). The theory seems to be supported by empirical data as shown in Schwartz et al. (1997a, 1997b) who examined 317 languages from the UCLA Phonological Segment Inventory Database (UPSID). Many dialectal studies conducted the past 20 years or so, however, show that there do exist extensive differences in the acoustic characteristics of vowels across dialects resulting in asymmetrical production systems (e.g. Clopper et al. 2005, 2006; Hagiwara 1997; Jacewicz et al. 2006; Recasens 2006). Such findings support a less stringent version of the Theory of Adaptive Dispersion whereby vowels are sufficiently dispersed (Lindblom 1986, 1990).

Cross-dialectal research on vowel perception is limited as compared to research on production. Most studies have examined how dialectal experience affects dialect classification (Clopper & Pisoni 2004; Evans & Iverson 2004; Labov & Ash 1997; Sumner & Samuel 2009). With respect to Greek, previous work has examined vowel perception in Standard Modern Greek (SMG), showing that vowels are well separated in the perceptual space (Botinis et al. 1997; Haws & Fourakis 1995).

This study examined the perception of vowels in SMG and two regional dialects, Cretan Greek (a Southern dialect) and Kozani Greek (a Northern dialect). All three systems have five vowels /i, e, a, o, u/[1]. Data were collected as part of a larger project investigating the phonetics and phonology of vowels across Greek dialects (http://www.vocalect.eu). The perceptual space was calculated following a method-of-adjustment procedure (Evans & Iverson, 2007) whereby participants chose vowel best exemplars (prototypes) for the five Greek vowels in a 5-dimensional space that included F1 and F2 formant movement (i.e. onset and offset of the F1 and F2 formant frequencies) as well as duration[2]. The production studies mentioned above report cross-dialectal asymmetries in the acoustic vowel space but since the perceptual space is maximized compared to the production one (the 'hyperspace' effect, see Johnson 2000; Johnson et al. 1993), it is possible that the acoustic vowel space is less affected by dialect than the perceptual one. The study therefore aimed at eliciting cross-dialectal data on the perception of Greek vowels that are currently lacking in the literature and at comparing the perception of vowels by dialectal speakers of the same language (Greek).

2. Method

[1] In unstressed position, Kozani Greek high vowels /i/ and /u/ are deleted ('high vowel deletion') and mid vowels /e/ and /o/ are raised to /i/ and /u/ respectively ('mid vowel raising') (Topintzi & Baltazani 2012; Trudgill 2003; Lengeris et al. 2016; Dinas 2005).
[2] The duration of vowel best exemplars is not discussed in this paper.

In I. Kappa & M. Tzakosta (eds), 2019. *Proceedings of the 7th International Conference on Modern Greek Dialects and Linguistic Theory*, 126-132. Patras: University of Patras.

2.1. Participants

Thirty participants were tested (ten SMG, ten Cretan Greek and ten Kozani Greek speakers). Their mean age was 61 years (range = 49 - 73 years). The Cretan and the Kozani speakers all had distinctive regional accents. None reported any hearing or language impairment.

2.2. Stimuli

The perceptual stimuli consisted of a synthetic vowel embedded within a carrier sentence *Πες* /pVta/ *ξανά* 'say /pVta/ again' spoken by a native speaker of the participants' dialect (e.g. throughout the perception experiment, the SMG participants heard a carrier sentence spoken by an SMG speaker). The /pVta/ context (stressed on the first syllable) created a minimally contrastive set of three words (*πίτα* /pita/ 'pie', *πέτα* /peta/ 'fly', and *πάτα* /pata/ 'press') and two non-words (**πότα* /pota/ and **πούτα* /puta/). The non-words are phonotactically acceptable in Greek and their production was not expected to be affected in any way.

Before synthesizing the vowels, the sentences produced by the SMG, the Cretan and the Kozani speaker were normalized to a 'model' speaker in terms formant frequencies and median pitch using signal processing in PRAAT (Boersma & Weenink 2016). This was done to reduce any effect any vocal tract differences between the three speakers might have on participants' location of vowel best exemplars (see Evans and Iverson 2007 for details).

The synthetic vowels were created using a Klatt synthesizer (Klatt & Klatt 1990) in cascade configuration so that they matched the vowels spoken by the corresponding speaker in terms of F0 and amplitude. The rest of the synthesis parameters were kept the same across vowels and dialects. These parameters were the F4 and F5 frequencies (3500 and 4500 Hz respectively), the formant bandwidths (B1=100, B2=180, B3=250, B4=300, B5=550), the tilt (TL=0 dB slope) and the open quotient (OQ=60%). The F1 and F2 frequencies changed in a linear way from the beginning to the end of the vowel. F1 formant frequency ranged between 5 and 15 Equal Rectangular Bandwidth (ERB) (Glasberg & Moore 1990). F2 formant frequency started from 10 ERB, was at least 1 ERB higher than F1 and reached a limit that was defined by the equation F2 = 25 - (F1-5) /2. The synthesized vowels were 1 ERB apart from each other. For each dialect, 109,375 vowels were synthesized.

Figure 1: *Experiment interface*

Figure 2: *Experiment interface*

2.3. Procedure

Testing was done in quiet rooms using a laptop computer and high-quality headphones (DT 770 PRO) with the help of a research assistant who was a native speaker of the dialect tested each time. During the experiment, the participants saw on the screen a /pVta/ structure written both in Greek orthography and Roman alphabet (e.g. ***πίτα*** and **pita** respectively) and heard the carrier sentence with a synthetic vowel in /pVta/ (see Figure 1). They were asked to rate how close the vowel was to a good exemplar of the vowel displayed on the screen by clicking on a continuous bar (see Figure 2).

Based on participants' responses, a goodness optimization method (Evans & Iverson 2004, 2007; Iverson & Evans 2003, 2009) found best exemplar locations of the five Greek vowels. During testing, an algorithm searched along 7 vectors (straight-line paths cutting through the synthesized space). The best exemplar on each vector was found after 5 trials per vector. The whole process required just 35 trials for each vowel despite the large number of synthesized vowels available to listeners and was completed after 175 trials (5 vowels × 35 trials) in about half an hour. Vector 1 found an approximation of participants' best exemplar by passing through the location of the natural production of the target vowel and the middle of the vowel space (F1 = 500 Hz, F2 = 1500 Hz) without varying duration; Vector 2 varied duration only while all other parameters were fixed; Vector 3 only varied the F1 and F2 onset frequencies while Vector 4 was orthogonal to Vector 3; Vector 5 only varied the F1 and F2 offset frequencies while Vector 6 was orthogonal to Vector 5; Vector 7 simultaneously varied all dimensions allowing the participants to fine-tune their answers for each vowel.

3. Results

The mean best exemplars (in ERB) that portray the perceptual vowel space in SMG, Cretan Greek and Kozani Greek are shown in Figures 1-3 respectively. It can be seen that, across dialects, the five Greek vowels were well separated in the perceptual space with no overlapping between vowels. Having said that, there were differences between dialects in the precise positioning of vowels in the perceptual space, the distance between adjacent vowels, and the total perceptual space area covered. The perceptual vowel space in SMG was the most symmetrical one compared to the perceptual vowel spaces in Cretan Greek and Kozani Greek. This finding is consistent with dialectal studies examining the acoustic vowel spaces in Greek (Trudgill 2009) and other languages (e.g. Clopper et al. 2005). For example, in both Cretan and

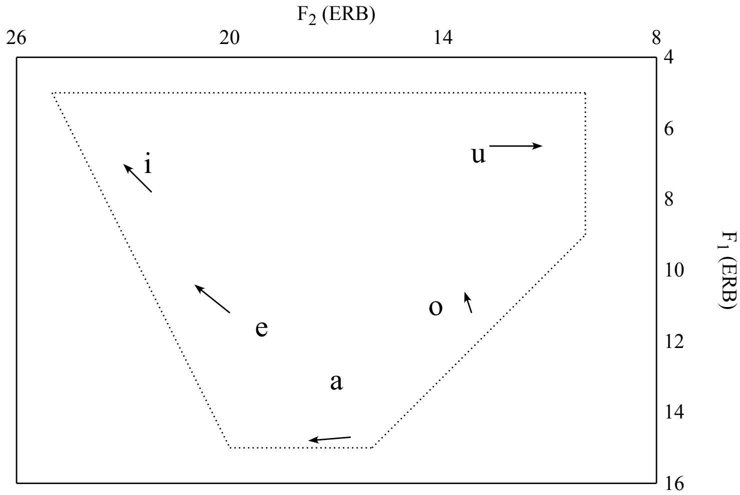

Figure 1: *Mean best exemplars (in ERB) of SMG vowels. Vowels are represented as arrows from the onset to the offset of the F1 and F2 formant frequencies. The limits of the synthesized vowels available to listeners are indicated by the dotted line.*

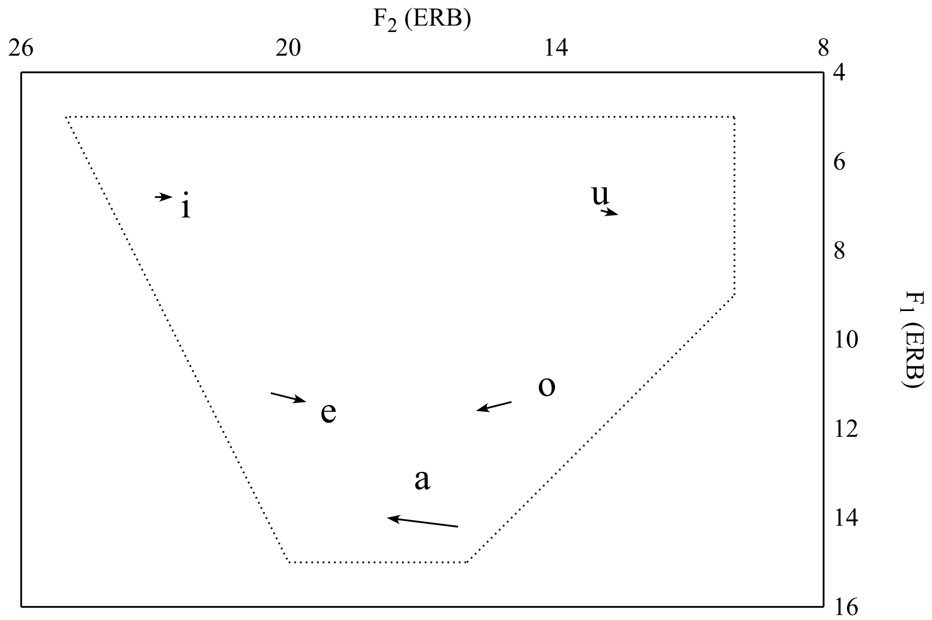

Figure 2: *Mean best exemplars (in ERB) of Cretan Greek vowels. Vowels are represented as arrows from the onset to the offset of the F1 and F2 formant frequencies. The limits of the synthesized vowels available to listeners are indicated by the dotted line.*

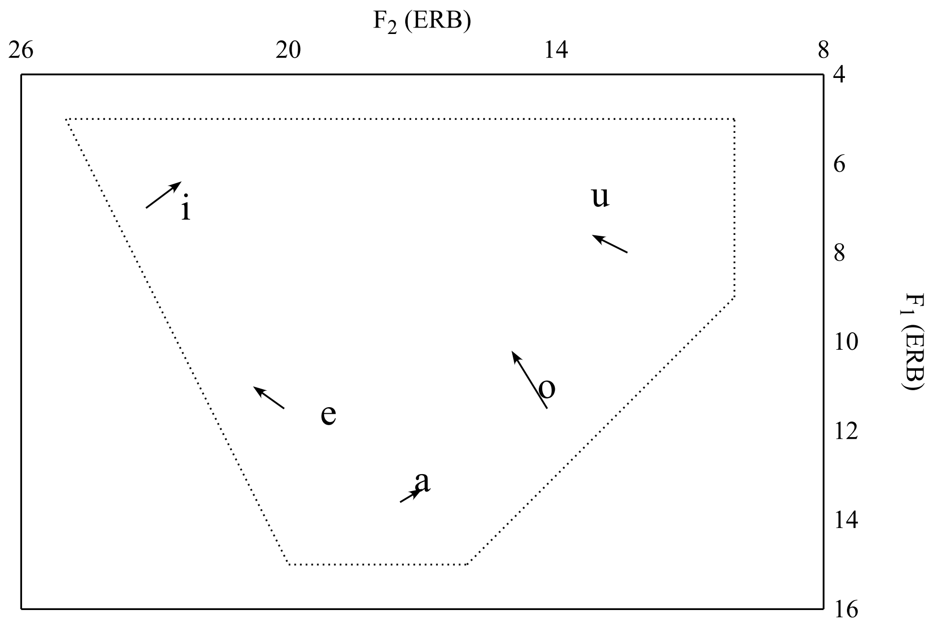

Figure 3: *Mean best exemplars (in ERB) of Kozani Greek vowels. Vowels are represented as arrows from the onset to the offset of the F1 and F2 formant frequencies. The limits of the synthesized vowels available to listeners are indicated by the dotted line.*

Kozani Greek the best exemplars for /e/ and /o/ were much closer to /a/ than to /i/ and /u/.

Table 1 displays the Euclidean distance between adjacent best exemplars (/i/-/e/, /e/-/a/, /a/-/o/, /o/-/u/, and /u/-/i/) in SMG, Cretan Greek and the Kozani Greek. To calculate the Euclidean distance, the four-dimensional coordinate for each best exemplar (i.e. onset and offset F1-F2 frequencies) was transformed to a two-dimensional coordinate by averaging the onset and offset of the F1 and F2 formant frequencies of each best exemplar, which removed the F1-F2 formant movement. A repeated-measures ANOVA was run on Euclidean distance with adjacent best exemplars (i.e. pair of best exemplars) (5 levels) and dialect (3 levels) as factors.

The ANOVA yielded a significant main effect of pair, $F(4,86) = 146.8$, $p <0.001$ and a significant pair × dialect interaction, $F(8,86) = 6.2$, $p <0.001$. This indicated that dialect did not have the same effect on Euclidean distance across pairs of best exemplars. Simple effect tests showed that dialect significantly affected the Euclidean distance in /i/-/e/ (Crete, Kozani > SMG), /a/-/o/ (SMG > Kozani > Crete) and /o/-/u/ (Crete > Kozani), $p <0.05$.

Apart from adjacent best exemplar distances, dialect also affected the total perceptual space areas. To compute vowel space areas, the Greek five-vowel space was divided into three triangles, the area of each triangle was calculated using

	Euclidean distance (ERB)				
	/i/-/e/	/e/-/a/	/a/-/o/	/o/-/u/	/u/-/i/
SMG	4.20	5.29	5.67	4.73	10.8
Cretan Greek	5.48	4.3	3.04	5.17	10.22
Kozani Greek	5.35	4.10	4.09	4.00	10.24

Table 1: *Euclidean distance (in ERB) between adjacent best exemplars in SMG, Cretan Greek and Kozani Greek.*

Heron's formula and the triangles were summed. The vowel space areas were 53.4 ERB^2 for SGM speakers, 38.9 ERB^2 for Cretan speakers and 40.2 ERB^2 for Kozani speakers. An analysis of variance (ANOVA) showed that the effect of dialect on vowel space areas was significant $F(2,29) = 6.7$, $p < 0.05$. Pairwise comparisons (Bonferroni adjusted) showed that the SMG vowel space area was the largest with no difference between the Cretan and the Kozani vowel space areas.

4. Discussion

This study compared the perceptual vowel spaces of SMG and two regional Greek dialects, Cretan Greek (a Southern dialect) and Kozani Greek (a Northern dialect), all with five vowels in their systems. Ten participants from each dialect selected best exemplars for vowels embedded in a carrier sentence uttered by a speaker of their dialect. A goodness optimization method allowed finding best exemplars from a very large stimulus set in less than half an hour.

The results showed that despite having well dispersed perceptual systems, SMG, Cretan Greek and Kozani Greek differed in terms of the exact positioning of vowels, the distances between adjacent vowels and their total space areas. The SMG system was the most symmetrical system compared to those of Cretan Greek and Kozani Greek. In addition, the SMG system covered the largest perceptual area compared to the other two. The two non-standard dialects therefore had less symmetrical and less expanded systems than the standard variety, a finding that is in agreement with cross-dialectal studies examining vowel production (e.g. Clopper et al. 2005; Trudgill 2009). It seems that while perceptual targets are generally more extreme than vowel production (Johnson 2000; Johnson et al. 1993), this perhaps reduces but does not completely remove dialect-specific tendencies in listeners' perception of vowels.

Acknowledgements

This research has been co-financed by the European Union (European Social Fund – ESF) and Greek national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework (NSRF)-Research Funding Program: THALES Investing in knowledge society through the European Social Fund, grant number MIS 379396.

5. References

Boersman, P., & Weenink, D. (2005) Praat: *Doing phonetics by computer [Computer software]*. University of Amsterdam, The Netherlands.

Botinis, A., Fourakis, M. & Hawks, J.W. (1997) A Perceptual Study of the Greek Vowel Space Using Synthetic Stimuli. *Proceedings of the 5th Eurospeech,* 1307-1310.

Clopper, C.G. & Pisoni, D.B. (2004) Some Acoustic Cues for the Perceptual Categorization of American English Regional Dialects. *Journal of Phonetics 32*: 111-140.

Clopper, C.G., & Pisoni, D.B. (2006) The Nationwide Speech Project: A New Corpus of American English Dialects. *Speech Communication 48*: 633-644.
Clopper, C., Pisoni, D. & de Jong, K. (2005) Acoustic Characteristics of the Vowel Systems of Six Regional Varities of American English. *Journal of the Acoustical Society of America 118*: 1661-1676.
Dinas K. (2005) *The dialect of Kozani*. Kozani: Institute of Book and Reading.
Evans, B. G. & Iverson, P. (2004) Vowel Normalization for Accent: An Investigation of Best Exemplar Locations in Northern and Southern British English Sentences. *Journal of the Acoustical Society of America 115(1)*: 352-361.
Evans, B.G. & Iverson, P. (2007) Plasticity in Vowel Perception and Production: A Study of Accent in Young Adults. *Journal of the Acoustical Society of America 121*: 3814-3826.
Glasberg, B.R. & Moore, B.C.J. (1990) Derivation of Auditory Filter Shapes From Notched-noise Data. *Hearing Research 7*: 103-138.
Hagiwara, R. (1997) Dialectal Variation and Formant Frequency: The American English Vowels Revisited. *Journal of the Acoustical Society of America 102(1)*: 655-658.
Haws, J.W. & Fourakis, M. (1995) The Perceptual Vowel Spaces of American English and Modern Greek: A Comparison. *Language and Speech 38*: 237-252.
Iverson, P. & Evans, B.G. (2003) A Goodness Optimization Method for Investigating Phonetic Categorization. *Proceedings of the 15th International Congress of Phonetic Sciences*. Barcelona, 2217-2220.
Iverson, P. & Evans, B.G. (2009) Learning English Vowels with Different First-language Vowel Systems II: Auditory Training for Native Spanish and German Speakers. *Journal of the Acoustical Society of America 126 (2)*: 866-877.
Jacewicz, E., Fox, R., Holt, Y. & Salmons, J. (2006) Acoustic Characteristics of Vowels in Three Regional Dialects of American English. *Journal of the Acoustical Society of America 120(5)*: 3294.
Johnson, K. (2000) Adaptive Dispersion in Vowel Perception. *Phonetica 57*: 181-188.
Johnson, K., Flemming, E. & Wright, R. (1993) The Hyperspace Effect: Phonetic Targets Are Hyperarticulated. *Language 69*: 505-528.
Klatt, D. H. & Klatt, L.C. (1990) Analysis, Synthesis, and Perception of Voice Quality Variations Among Female and Male Talkers. *Journal of the Acoustical Society of America 87*: 820– 857.
Labov, W. & Ash, S. (1997) Understanding Birmingham. In: C. Bernstein, T. Nunnally & R. Sabino (Eds.), *Language variety in the South revisited*. AL: Alabama University Press, 508-573.
Liljencrants, J. & Lindblom, B. (1972) Numerical Simulations of Vowel Quality Systems: The Role of Perceptual Contrast. *Language 48*: 839-862.
Lindblom, B. (1986) Phonetic Universals in Vowel Systems. In: J.J. Ohala, J.J. Jaeger (Eds.), *Experimental phonology*. New York: Academic Press, 13-44.
Lindblom, B. (1990) On the Notion of Possible Speech Sound. *Journal of Phonetics 18*: 135-152.
Lengeris, A., Kainada, E. & Topintzi, N. (2016) Vowel Raising, Deletion and Diphthongization in Kozani Greek. *Proceedings of the 6th International Conference on Modern Greek Dialects & Linguistic Theory*. Patras, 93-101.
Recasens, D. & Espinosa, A. (2006) Dispersion and Variability in Catalan Vowels. *Speech Communication 48*: 645-666.
Schwartz J.-L., Louis J.-B, Nathalie V. & Christian, A. (1997a) Major Trends in Vowel System Inventories. *Journal of Phonetics 25(3)*: 233-253.
Schwartz J.-L., Louis J.-B, Nathalie V. & Christian, A. (1997b) The Dispersion- Focalization Theory of Vowel Systems. *Journal of Phonetics 25(3)*: 255-286.
Sumner, M. & Samuel, A.G. (2009) The Role of Experience in the Processing of Cross-Dialectal Variation. *Journal of Memory and Language 60*: 487 -501.

Topintzi, N. & Baltazani, M. (2012) The Acoustics of High-vowel Loss in a Northern Greek Dialect and Typological Implications. In: P. Hoole, L. Bombien, M. Pouplier, Ch. Mooshammer & B. Kühnert. (Eds.), *Consonant clusters and structural complexity*. Berlin: Mouton de Gruyter, 373-402.

Trudgill, P. (2003) Modern Greek Dialects: a Preliminary Classification. *Journal of Greek linguistics 4*: 45-64.

Trudgill, P. (2009) Greek Dialect Vowel Systems, Vowel Dispersion Theory, and Sociolinguistic Typology. *Journal of Greek Linguistics 9*: 165-182.