Perception of fricative voice distinctions in Greek

Elina Nirgianaki[1], Antonis Botinis[1] & Marios Fourakis[2]
[1]Laboratory of Phonetics and Computational Linguistics, University of Athens, Greece
[2]Department of Communication Sciences and Disorders, University of Wisconsin-Madison, USA

Abstract

This is a study of the perception of the voice distinction in Greek fricatives across different places of articulation. The results indicate that: (1) the acoustic correlate of voicing is not a categorical perception effect of fricative voice distinctions; at least not across all places of articulation, and 2) noise duration has a significant effect on fricative voice perception across all places of articulation, but only in conjunction with glottal vibration.

Introduction

Voice contrasts in Greek fricatives are present at all places of articulation, i.e. labial (labiodental) [f~v], dental [θ~ð], alveolar [s~z], palatal [ç~ʝ] and velar [x~ɣ]. Palatal fricatives have normally no phonemic status but are as a rule context-dependent (Botinis, 2011).

These contrasts affect three main acoustic characteristics. First, voiced fricatives are fully voiced whereas voiceless ones are typically voiceless. Second, voiceless fricatives are considerable longer than voiced ones (see *Table 1*). Third, [voice] differences trigger compensatory durational adjustments on the following vowel: vowel durations are longer after voiced and shorter after voiceless fricatives (Nirgianaki, 2009, 2010, 2013).

In the present study, an experiment was carried out to determine the perceptual correlates of the voicing contrast in Greek fricatives. Two main questions were addressed: (1) does the presence of voicing induce a categorical perceptual effect in distinguishing voiced from voiceless fricatives? And (2), what is the effect of different durational patterns on the perception of the [voice] contrast in these fricatives? Although the acoustics of fricatives (for both place of articulation and [voice]) have been studied extensively (e.g. Baum & Blumstein, 1987; Crystal & House, 1988; Jongman et al., 2000; Maniwa, Jongman & Wade, 2009), their effects on perception have not drawn particular attention. On the other hand, the [voice] contrast in fricatives occupies a central place in the phonetic system of several different language families.

Experimental methodology

Ten (10) female listeners, 20–35 years old, participated to the experiment. All were native speakers of Greek, speaking what is commonly called standard Athenian. None of them had any history of speech or hearing disorders and none had recorded the experimental material. All speakers participated in the experiments as volunteers.

The speech material was recorded by a female native speaker. The six (6) Greek fricative consonants [f], [v], [s], [z], [x] and [ɣ] were recorded in real, two-syllable words of the structure CVCV, stressed on the first syllable. The vowel [a] followed the first fricative consonant. Five repetitions were recorded, from which one was used for the perception experiments. The words were placed in the carrier phrase [ˈipa … ksaˈna] (I said … again). The duration of each fricative and the following vowel were measured with the simultaneous consultation of the waveform and the wideband spectrogram and the mean duration was calculated (*Table 1*).

Table 1. Mean durations (in ms) of each fricative and the following vowel (/a/).

Word	Fricative	Vowel
ˈfata	91	102
ˈvata	69	117
ˈsali	125	139
ˈzali	91	155
ˈxamo	103	116
ˈɣamo	80	140

Then, three extra files were created in Praat (Boersma & Weenink, 2009) for each file that contained a voiced fricative: (a). a file with the voiced fricative prolonged (named ‘Long C’ – each voiced fricative was prolonged by adding to it the mean percentage difference that it had from its voiceless counterpart) (b). a file with the vowel followed the voiced fricative shortened (named ‘Short V’ – the following vowel was shortened by removing the mean percentage difference it had from the one that followed its voiceless counterpart), and (c). a file with the voiced fricative having no voice bar (named 'No voicing' – for each voiced fricative, the frequencies between 0–500 Hz were filtered out). Four more files were created for each voiced fricative, which resulted from all possible combinations of these three files (i.e. 'Long C + short V', 'Long C + no voicing', 'Short V + no voicing', 'Long C + short V + no voicing').

The stimuli used for the perception experiment were the files that contained a word with a voiced fricative (original, manipulated or a combination of manipulated files). Each such file was heard ten (10) times, yielding a total of eighty (80) stimuli per fricative and 240 stimuli in total.

Three experimental sets were conducted for each participant in a quiet room at the University of Athens. All these experimental sets were two-alternative forced-choice identification tasks, for which clear instructions were provided in Greek. Participants were informed that they would hear the phrase [ˈipa … ksaˈna] "I said … again" containing one of the two words [ˈfata] or [ˈvata] for the first experimental set, [ˈsali] or [ˈzali] for the second experimental set, and [ˈxamo] or [ˈɣamo] for the third experimental set. Participating listeners were asked to identify the word by pressing the appropriate button.

Stimuli were presented to listeners via a program created on Matlab (Mathworks, 2011) on a Hewlett Packard laptop computer over Direct Sound headphones. Following a short familiarization phase of three items, the stimuli were played in random order once each, after pressing a "Play Next" button appearing on the screen.

After each stimulus was played, a prompt appeared on screen containing the two possible response buttons. Participants could press only one of them and after pressing it, the prompt disappeared and they should press the "Play Next" button in order to hear the next item. The two response buttons contained the two Greek words written in Greek orthography, as follows: "φάτα" ([ˈfata]), "βάτα" ([ˈvata]) for the first experimental set, "σάλι" ([ˈsali]), "ζάλη" ([ˈzali]) for the second one, and "χάμω" ([ˈxamo]), "γάμο" ([ˈɣamo]) for the third one. Except for the two buttons corresponding to the two words, a "Replay" button was also included in the prompt, which if pressed, the same phrase was played again.

Listeners could hear each sound three times maximum, by pressing this button two times at most. For each participant, all three experimental sets lasted approximately ¾ hour, including a short break (of around 5 minutes) after each one.

Results

The results of the perception experiment are presented in the following *Table 2* and *Figures 1–2*. Statistical analysis was carried out using the SPSS 19.0 (SPSS Inc., 2010) software package.

It was revealed that voiced fricatives were identified mainly as voiceless, when they were lengthened and their voicing had been eliminated (conditions 'Long C + no voicing' and 'Long C + short V + no voicing'). However, there was a systematic effect of place of articulation on listeners' performance. We used the Generalized Estimating Equations (GEE) analysis, in order to conduct the equivalent of a repeated measures ANOVA. The logistic linking function was used and each of the ten repetitions was treated as a repeated measurement. Place of articulation and condition were within-subjects factors; participant was a between-subjects factor.

Participant had a highly significant main effect (Wald χ^2 (9) = 215360.507, $p < .0001$): the identification of voiceless sounds ranged from 16.3% to 49.6% (*Figure 1*). Condition was also significant (Wald χ^2 (7) = 1982.839, $p < .0001$). The identification of voiceless sounds ranged from 2.7% in the 'Short V' condition to 73% in the 'Long C + no voicing' condition (*Table 2*).

There was a significant main effect of place of articulation (Wald χ^2 (2) = 787.771, $p < .0001$) with higher identification of voiceless sounds in the velar place (velars: 50.2%, alveolars: 31.6%, labiodentals: 14.5%). Individual comparisons against the velars (which showed the highest performance on voiceless identification) showed that in the other two places of articulation voiceless sounds were identified significantly less (labiodentals: Wald χ^2 (1) = 79.683, $p < .0001$; alveolars: Wald χ^2 (1) = 75.290, $p < .0001$).

A place x condition interaction (Wald χ^2 (9) = 209.027, $p < .0001$) was mainly driven by the velars (*Figure 2*), which, contrary to the other two places, were identified as voiceless in all 'no voicing' conditions (i.e. apart from 'Long C + no voicing' and 'Long C + short V + no voicing', also in 'No voicing' and 'Short V + no voicing' conditions).

Table 2. Mean percentage score of voiceless response (%) for each place of articulation and averaged across all three places.

CONDITION	Labials	Alveolars	Velars	Mean
Original file	0	0	9	3
Long C	0	29	20	16.3
Short V	0	1	7	2.7
Long C + short V	0	27	21	16
No voicing	13	17	71	33.7
Short V + no voicing	21	22	78	40.3
Long C + no voicing	42	80	97	73
Long C + short V + no voicing	40	77	99	72

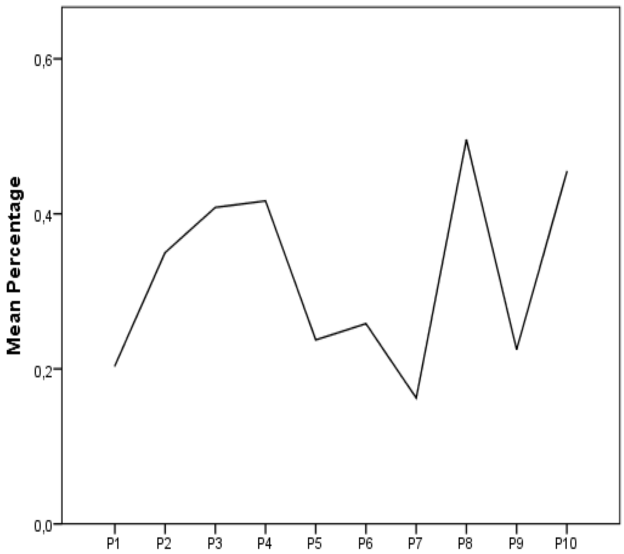

Figure 1. Mean percentage score (%) of voiceless response for each participant.

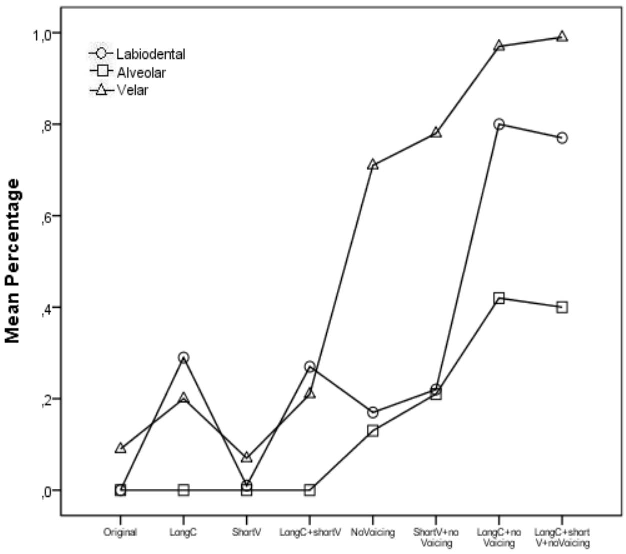

Figure 2. Mean percentage score (%) of voiceless response as a function of place of articulation and condition.

Discussion and conclusions

In accordance with the questions put in the Introduction, the results of this study indicate the following. First, Greek voiced fricatives of all the examined places of articulation were identified mostly as voiceless when they were prolonged to match the duration of their voiceless counterparts and had their glottal vibration removed. Second, voiced velar fricatives were identified as voiceless significantly more than alveolars and labiodentals (velars > alveolars > labiodentals). Third, velar fricatives were also identified as voiceless when they had just their glottal vibration removed.

These results have several implications: 1) the acoustic correlate of voicing is not a categorical perception effect of fricative voice distinctions; at least not across all places of articulation, 2) although the duration of the post-fricative vowel acts compensatory acoustically, it does not account for the perception of the fricative's voicing status, and 3) noise duration has a significant effect on fricative voice perception, but only in conjunction with glottal vibration.

Similarly, the decrease of the frication noise duration of English voiceless fricatives (in syllable-initial position) has been shown to result in an increase of their perception as their voiced counterparts (Cole & Cooper, 1975). Moreover, the fact that there are durational differences between fricatives of the same voicing but different places of articulation (also reported for fricatives with different amplitude by Behrens and Blumstein, 1988) does not seem to play any role, since voiced velars were identified the most as voiceless ones, although shorter than alveolars, and with smaller percentage of lengthening (velars: 28.7%, alveolars: 37.7%, labiodentals: 31.2%).

However, the present study did not examine the effect of F1 (both in terms of onset frequency and change) at the CV boundary for fricatives of different places of articulation, which may have played an important role on listeners' voicing judgements. Results of Stevens et al., (1992) suggest that "listeners base their voicing judgments of intervocalic fricatives on an assessment of the time interval in the fricative during which there is no glottal vibration, and this time interval depends on the extent to which the F1 transitions are truncated at the consonant boundaries".

As stated in the Introduction, [voice] contrasts occur at all places of articulation in Greek. The contrast is a binary specification at the underlying representation level, i.e. ±[voice]. However, in accordance with the results of the present study, the functional binary nature of this contrast is related to several acoustic correlates, in addition to voicing, including, in the first place, duration. Further acoustic correlates include the durations of post-fricative vowels and, presumably, the acoustic structure of the immediate syllabic context in general.

Earlier research indicates that duration is a fairly constant acoustic correlate of voicing distinctions in Greek (Nirgianaki, 2009, 2010, 2013). However, this is evident in simple CV syllabic contexts. In palatalised contexts, which are very common in Greek, duration does not seem to contribute to voicing fricative distinctions. Thus, in words such as [kuˈpça] (oars) and [kuˈbʝa] (buttons) the duration of the voiceless and voiced fricative after the stop is approximately the same. This surface phonetic representation has been derived from a series of rules that result in a voiceless and voiced fricative production after voiceless and voiced consonant productions respectively (Botinis, 2011). In other words, the fricative voicing distinction in a palatalised context is related to the immediate syllabic context, which carries the respective voicing distinctions.

Thus, the duration patterns corresponding to fricative [voice] contrasts in different contexts is not related to any inherent production specifications but rather to the overall sequence structure. Given that there is no phonemic length distinction in Greek consonants, it is remarkable that durational patterns seem to contribute to a variety of surface phonetic distinctions in the consonant system of Greek. On the other hand, an effective speech communication process has to guarantee that the surface phonetic contrast be effectively conveyed and this results in the association of fricative [voice] distinctions with several acoustic correlates.

Acknowledgements

Thanks to Special Research Account of Athens University for a travel grant to Fonetik 2013 conference.

References

Baum, S. R. & S. E. Blumstein. 1987. Preliminary observations on the use of duration as a cue to syllable-initial fricative consonant voicing in English. *Journal of the Acoustical Society of America* 82:1073–1077.

Behrens, S. J. & S. E. Blumstein. 1988. Acoustic characteristics of English voiceless fricatives: A descriptive analysis. *J. Phonet.*16, 295-298.

Boersma, P. & D. Weenink. 2009. Praat: doing phonetics by computer, version 5.1.01. www.praat.org

Botinis, A. 2011. *The Phonetics of Greek* (in Greek, Fonitiki tis Elinikis). Athens: ISEL Editions.

Cole, R.A. & W. E. Cooper. 1975. Perception of voicing in English affricates and fricatives *Journal of the Acoustical Society of America* 58:1280–1287.

Crystal, T. & A. House. 1988. Segmental durations in connected speech signals: Current results. *Journal of the Acoustical Society of America* 83:1553–1573.

Jongman, A., Wayland, R. & S. Wong. 2000. Acoustic characteristics of English fricatives. *Journal of the Acoustical Society of America* 108:1252–1263.

Maniwa, K., A. Jongman & T. Wade. 2009. Acoustic characteristics of clearly spoken English fricatives. *Journal of the Acoustical Society of America* 125:3962–3973.

Mathworks, Inc., The. 2000. MATLAB R2011b, The language of technical computing, version 7.13.0.564.

Nirgianaki, E. 2013. *Acoustic and Perceptual Characteristics of Greek Fricative Consonants* (in Greek). PhD Thesis, University of Athens.

Nirgianaki, E., A. Chaida & M. Fourakis. 2010. Acoustic structure of fricative consonants in Greek. *Proceedings ExLing 2010*, Athens, Greece, 125–128.

Nirgianaki, E., A. Chaida & M. Fourakis. 2009. Temporal characteristics of Greek fricatives. *Proc. ICGL*, 25–33.

SPSS Inc. 2010. SPSS for Windows: version 9 2010 SPSS Chicago IL.

Stevens, K. N., S. E. Blumstein, L. Glicksman, M. Burton & K. Kurowski. 1992. Acoustic and perceptual characteristics of voicing in fricatives and fricative clusters. *Journal of the Acoustical Society of America* 91:2979-3000.

Proceedings of Fonetik 2013

The XXVIth Annual Phonetics Meeting
12–13 June 2013, Linköping University
Linköping, Sweden

Studies in Language and Culture
no. 21

Robert Eklund, editor

Linköping University

Conference website: www.liu.se/ikk/fonetik2013
Proceedings also available at: http://roberteklund.info/conferences/fonetik2013

Cover design and photographs by Robert Eklund
Photo of Claes-Christian Elert taken by Eva Strangert on the occasion of his 80th birthday

Proceedings of Fonetik 2013, the XXVIth Swedish Phonetics Conference
held at Linköping University, 12–13 June 2013
Studies in Language and Culture, no. 21
Editor: Robert Eklund
Department of Culture and Communication
Linköping University
SE-581 83 Linköping, Sweden

ISBN 978-91-7519-582-7
eISBN 978-91-7519-579-7
ISSN 1403-2570

Printed by LiU-Tryck, Linköping, Sweden, 2013