RECOGNIZING CHANGES IN HEBREW VOWEL HEIGHT AND VOWEL PLACE USING ONE FORMANT ONLY

Liat Kishon-Rabin, Noam Amir, Vera Bein & Tzipi Venezian

Department of Communication Disorders, Sackler Faculty of Medicine, Tel Aviv University, Israel
lrabin@post.tau.ac.il; noama@post.tau.ac.il

ABSTRACT

The purpose of this study was to determine whether a change of one formant only is sufficient to shift the recognition of one vowel category to another. Two sets of vowel continua were constructed from a naturally produced /o/. In one continuum, F2 was varied to shift from /o/ to /e/, and in the other F1 was varied to shift from /o/ to /u/. Identification curves were then collected from 20 normal-hearing female subjects. Results indicated that change in F2 was sufficient for all participants to shift from /o/ to /e/. In contrast, F1 was sufficient to shift from /o/ to /u/ for only half the participants. When a shift in perception occurred it was categorical in nature. These data suggest that in Hebrew, listeners probably use additional/other cues to perceive the vowel height contrast. Implications for vowel perception in special populations such as hearing-impaired with cochlear implants are discussed.

Keywords: vowels, formants, categorical perception

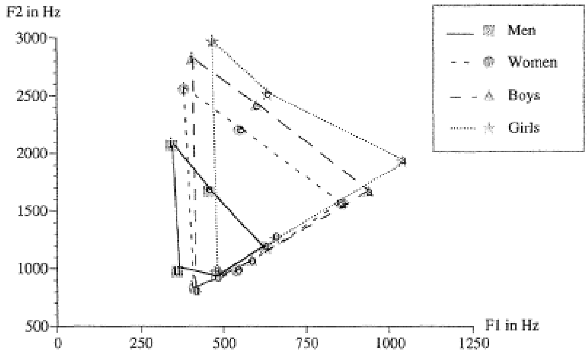
1. INTRODUCTION

Traditionally, the first two resonant frequencies (also known as formants) of the vocal tract, F1 and F2, are considered the main cues for the perception of vowels [4, 11]. The values of these first two formants are first and foremost influenced by the length of the vocal tract (following the source-filter model) and by the height and position of the tongue in the oral cavity where some narrowing of the vocal tract occurs for producing the different vowels. It is also suggested that F1 is typically influenced by the height of the tongue and therefore is used to cue vowels such as /o/ vs /u/, whereas F2 is influenced by the position of the tongue (front back) such as in vowels (/i/ vs /u/) [4, 12].

In the 220 recorded languages, there are a total of 60 vowel systems, of which 55 include the five extreme vowels /i,e,a,o,u/. The formant values of the vowels, including the extreme ones, vary between languages, dialects, speakers, gender and age groups [4]. Despite the large variability in formant values, listeners within a given language appear to identify the vowels of their language very well. Vowels have relatively high intensity and long duration, and are therefore also easy to recognize by hearing-impaired individuals with hearing aids [2].

There are recent data to suggest that prelingual hearing-impaired children with cochlear implants (CI) are well able to recognize vowels that differ by changes of F1 as well vowels that differ by changes of F2. In contrast, postlingual adults with CI had difficulty in recognizing vowels that differ by changes of F1 only. The fact that vowels cued by changes of F2 were easily perceived by the two groups of CI users can be explained by the fact that the changes of F2 were large enough to electrically stimulate two different places on the basilar membrane. However, changes in F1 were difficult to perceive by the CI users because they often stimulate the same electrode or two adjacent electrodes [7]. But if it is difficult, how is it that the prelingual children with CI were able to perform this task successfully? In a first attempt to answer this question, we wanted to determine whether F1 and F2 are indeed significant cues for the perception of vowel height and vowel place contrasts in Hebrew, a language that includes only the five extreme vowels. The Hebrew vowel space for children and adults is presented in Figure 1 [9].

Figure 1: The Hebrew vowel space for Men, Women, Boys and Girls, following Most, et al. [9].



Thus, the purpose of the present study was to evaluate in normally hearing native Hebrew speakers whether changes of F1 only were sufficient to signal a shift of vowel category from /o/ to /u/, and whether changes of F2 were sufficient to signal a shift of vowel category from /o/ to /e/. The data provided us with first time evidence as to the acoustic information necessary for recognizing vowels in Hebrew. It also provided us with possible explanations to the results of the CI suggesting that both the normally hearing and cochlear implantees use, in addition to F1, other acoustic cues that enable them to recognize the vowel height contrast.

2. METHOD

2.1. Participants

Twenty women participated in the experiment, 23-29 years of age (*M*=25.1). All were native speakers of Hebrew without previous experience in psychoacoustic experiments. All participants had normal hearing thresholds [1].

2.2. Stimuli

Stimuli consisted of two continua of speech tokens: one varied in F1 only, and the other in F2 only. The basis for these continua was a naturally produced /o/ by a male native Hebrew speaker. Recording was conducted in a sound treated room, using an AT892 head mounted microphone at a sampling rate of 48 kHz and 16 bits/sample. The first sequence morphed from /o/ to /e/, through changing of the second formant only, in nine steps of 122Hz, from the original 928 Hz to 1904 Hz. The second sequence morphed form /o/ to /u/ through changing of the first formant only, in steps of 16.75 Hz, from the original 469 Hz to 335 Hz. All manipulations were performed using the Klatt synthesis option in Praat. Stimuli were trimmed to duration of 150 ms.

Note that the final values of each vowel continuum were determined by averaging eight productions of the vowels /u/ and /e/. The formant values in all utterances were analyzed using Praat software, giving the mean values and STD shown in Table 1. Also shown for comparison are average formant values of a larger group of Hebrew male speakers (*N*=30) [9].

It can be seen that the exemplars of the vowels used in the present study have similar formant values to those measured in Most, et al. [9]. It can also be seen that F1 values for /o/ and /e/ are similar, and F2 values for /o/ and /u/ are similar.

Table 1: Mean formant values (in Hz) of the naturally produced Hebrew vowels (STD) in the present study compared to that of Most et al (200X).

		/o/	/u/	/e/
F1	Present study	469 (*55*)	335 (*23*)	422 (*57*)
	Most et al.	478 (46)	359 (31)	455 (40)
F2	Present study	928 (*78*)	932 (*183*)	1904 (*66*)
	Most et al.	944 (105)	979 (91)	1662 (171)

Each participant performed two listening tests, one on each of the above sequences. In the o-e test, they heard the stimuli morphing from /o/ to /e/, and in the o-u test, they heard the stimuli morphing from /o/ to /u/. Both listening tests were of the Two Alternative Forced Choice type: participants heard a stimulus and were asked to classify it into one of two possible vowel categories. In each test, each of the 9 stimuli was presented 10 times, giving a total of 90 presentations per test, all in random order.

Participants were randomly divided into two groups. One group performed the o-e test first, followed by the o-u experiment. The second group performed the tests in reverse order. The tests were conducted in an acoustically treated booth. All stimuli were presented at 60 dB HL through an audiometer, using TDH-50 earphones.

3. RESULTS

The responses of each participant to each pair of vowels were averaged and plotted for each vowel pair separately as a function of the shift in F1 or F2. These graphs, referred to as classification plots, were divided into four general classes. These ranged from class A, demonstrating classical categorical perception, where there is an abrupt shift in identification from one category to another, through plots representing some uncertainty, to the extreme case of no shift at all from one category to another (class D). Examples of these four types are presented in figure 2(A) through 2(D).

For each test, /o-e/ and /o-u/, the classification plots of each subject were examined and assigned to one of the above classes (A to D). Figure 3 shows the distribution of the plots among these classes, for the two tests. From figure 3 it is evident that changes in F2 are enough to change the perception of vowels from /o/ to /e/.

Interestingly, perception of /o/ vs. /u/ was less conclusive. On the one hand, only 60% of the listeners perceived a change from /o/ to /u/ based on F1 information only, whereas 30% perceived no transition at all from one vowel to the other. These speakers perceived all stimuli as /o/.

Figure 2: Four types of generic classification plots. (A) classical categorical perception; (B) categorical perception with a region of uncertainty; (C) perception of one category with uncertainty in the other; (D) no shift in perception.

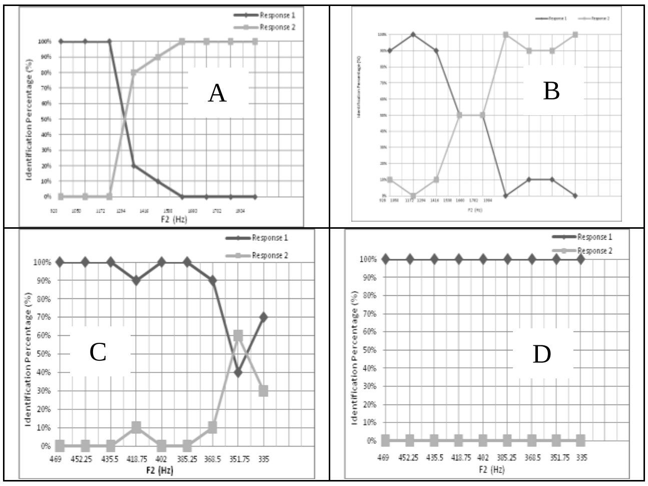
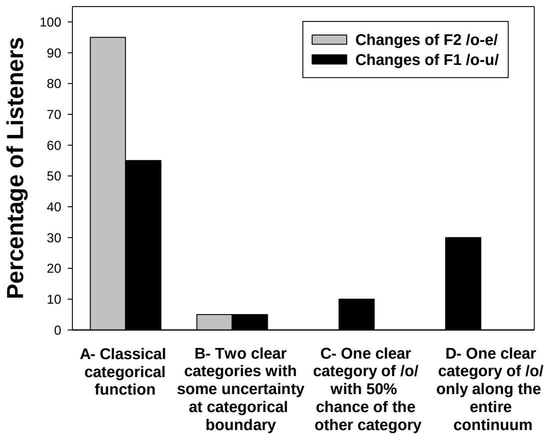


Figure 3: Distribution of classification plots of each vowel pair.



In the cases where a categorical boundary was found, the 50% crossover point was calculated for each listener and for each pair of vowels. Note that for class A graphs it was determined by linear interpolation, as the frequency where identification crosses the 50% point. For class B plots, where there is a region of uncertainty, the boundary frequency was taken as the lowest frequency at which the 50% point was crossed. Figures 4 and 5 show the distribution of boundary frequencies over the different listeners, for the /o-e/ and /o-u/ continua.

For the o-e continuum the average boundary frequency of F2 was 1341 Hz (*SD*=112, *Min*=1172, *Max*= 1558), which is roughly midway between the two extreme values.

Figure 4: Distribution of boundary frequencies for the o-e task. Dotted line in the middle is the mean boundary frequency.

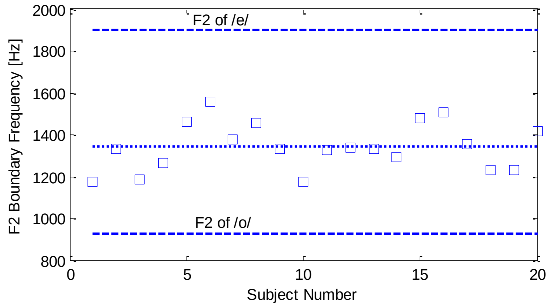
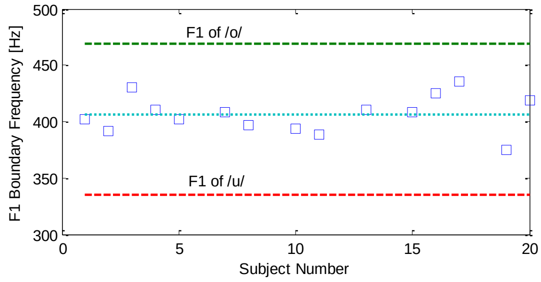


Figure 5: Distribution of boundary frequencies for the o-u task. Dotted line in the middle is the mean boundary frequency.



For the o-u continuum the average boundary frequency was 406 Hz (*SD*=16.5, *Min*=375, *Max*= 435), which again is roughly midway between the two extreme values. In this case there are missing points, since for some speakers there was no shift in the perceived category. It can be seen that for both continua there was relatively small inter-subject variability in boundary frequencies.

4. DISCUSSION

The main outcomes of this study were: 1) shifting a single formant *can* bring a shift in perception from one vowel to another in a categorical manner; 2) **F2** was found to be a sufficient acoustic cue for shifting between the perception of /o/ and /e/ for **all** listeners; 3) **F1** was found to be a sufficient acoustic cue for shifting between perception of /o/ and /u/ for approximately **half** of the listeners; 4) most listeners demonstrate a shift from one category to the other at a similar frequency value.

Our first finding that a continuous change of one acoustic parameter of a vowel resulted in clear steep identification functions is in contrast to previous findings of non-categorical perception of vowels [5, 13]. The hypothesis of previous researchers was that categorical perception requires a point of discontinuity in the vocal tract, such as in the case of plosives but not in vowels. Our findings suggest that such a discontinuity is not necessary. This has been found in suprasegmental studies as well, such as identification of question vs. statement [10]. The discrepancies between the present study and previous ones may be related to methodological factors, such as the fact that we used naturally produced stimuli as the basis for re-synthesis, whereas most other studies used synthetic stimuli. Another possible explanation may be related to the difference in vowel space and vowel inventory in English compared to Hebrew. Hebrew has 5 vowels, compared to 12 basic vowels of American English, not including diphthongs [8] and its vowel space is more triangular than the English one [9].

Our findings that changes of F1 and F2 alone were sufficient to shift the perception of one vowel to another are interesting in light of the notion that listeners rely on multiple cues for perception of vowels [4]. Moreover, changes in F2 were more effective in bringing about a shift in perception than changes in F1. One possible explanation is the difference in the range of the F2 continuum as compared to the F1 continuum: a ratio of 7.2 in Hz, and 3.6 in Barks. This clearly requires further investigation.

In summary, the results suggest that changes in in F2 are sufficient to indicate shifts in the position of the tongue from back to front (/o/ to /e/) in the Hebrew vowel system, whereas changes in F1 are sufficient for only some listeners to identify shifts in the height of the tongue from low to high (/o/ to /u/). Further research must be conducted to determine whether this is true for other pairs of vowels that differ in one significant phonological contrast in Hebrew and in other languages as well. Our data also suggest that listeners of the Hebrew language probably rely on cues other or in addition to F1 to perceive vowels such as /o/ compared to /u/. Such cues may be accessible via the cochlear implant device and perceived by the developing hearing-impaired child with a CI thus explaining the data of the CI. Why the postlingual adults with CI weren’t able to perceive these cues may be related to the fact that their brain was pre-wired to the acoustic information they used while they were hearing. Therefore lack of plasticity did not allow them to adjust to new information provided by the CI device, an advantage that the young prelingual hearing-impaired children had. These issues are currently under investigation.

5. REFERENCES

[1] American National Standards Institute. 1989. *ANSI S3.6-1989 - American National Standard Specification for Audiometers*. New York: American Institute of Physics.

[2] Boothroyd, A. 1984. Auditory perception of speech by subjects with sensirineural hearing loss. *Journal of Speech and Hearing Research* 27,134-144.

[3] Boothroyd, A. 1986. *Speech Acoustics and Perception* Austin: Pro-Ed Publishers.

[4] Borden, J.B, Harris, K.S., Raphael, J. 1994. *Speech Science Primer*. Baltimore: William and Wilkins.

[5] Fry, D.B., Abramson, A.S., Eimas, P.D., Liberman, A.M. 1962. The identification and discrimination of synthetic vowels. *Language and Speech* 5, 171-189.

[6] Kishon-Rabin, L., Rotshtein, S., Taitelbaum, R. 2002. Underlying mechanism for categorical perception: Tone-onset time and voice-onset time evidence of Hebrew voicing. *Journal of Basic Clinical Physiology and Pharmacology* 13(2), 117-134.

[7] Kishon-Rabin, L., Taitelbaum, R., Muchnik, C., Gehtler, I., Kronnenberg, J., Hildesheimer, M. 2002. Development of speech perception and production in children with cochlear implants. *Annals of Otology, Rhinology, and Laryngology* 111, 85-90.

[8] Kishon-Rabin, L., Taitelbaum, R., Tobin, Y., Hildesheimer, M. 1999. The effect of partially restored hearing on speech production of postlingually deafened adults with cochlear implants. *Journal of the Acoustical Society of America* 106(5), 2843-2857.

[9] Most, T., Amir, O., Tobin, Y. 2000. The Hebrew vowel system: raw and normalized acoustic data. *Language and Speech* 43(3), 295-308.

[10] Peng, S., Lu, N., Chatterjee, M. 2009. Effects of cooperating and conflicting cues on speech intonation recognition by cochlear implant users and normal hearing listeners. *Audiology & Neurotology* 14, 327-337.

[11] Peterson, G.E., Barney, H.L. 1952. Control methods used in a study of the vowels, *Journal of the Acoustical Society of America* 24, 175-184.

[12] Pols, L.C., Van der Kamp, L.J., Plomp, R. 1969. Perceptual and physical space of vowel sounds. *Journal of the Acoustical Society of America* 46, 458-467.

[13] Stevens, K., Liberman, A.M., Studdert-Kennedy, M., Ohman, S.E.G. 1969. Cross-language study of vowel perception. *Language and Speech* 12, 1-23.