

The role of formant amplitude in the perception of /i/ and /u/ in normal hearing listeners

Lacey Marshall, Teresa Enright, and Michael Kieffe

School of Human Communication Disorders, Dalhousie University, 5599 Fenwick Street, Halifax, NS, Canada, B3H 1R2

1. INTRODUCTION

It is known that each vowel in English has a unique configuration of formants, and this configuration is thought to help in the discrimination of vowel sounds. It is generally accepted that the three lowest formants are the most important for vowel perception (Rosner & Pickering, 1994). However, it is still debated whether it is formant frequency, spectral shape, or some combination of formant frequency, spectral shape, and formant amplitude that makes the largest contribution to vowel perception.

Although early studies indicated that formant amplitude plays only a minor role in vowel perception (e.g., Miller, 1984), recent research has indicated that formant amplitude may be more important than previously thought (Nabelek et al., 1992). However, effects of formant amplitude as observed in Nabelek et al. (1992) could instead be due to relatively basic psychoacoustic phenomena such as simultaneous masking. Perhaps related to this, Ainsworth and Millar (1972) concluded that listeners' perception of vowels changed when the amplitude of F_2 was decreased, but only when the difference in the amplitudes of F_1 and F_2 was no more than 28 dB.

Aaltonen (1985) performed a series of perception experiments in which the amplitude of F_2 and F_3 was varied in the Finnish vowels /i/ and /y/. It was found that reducing the amplitude of F_2 alone increased the number of /i/ responses, while reducing the amplitude of F_3 alone resulted in more /y/ responses. Aaltonen suggested that the reason for this response pattern was that when the amplitude of F_2 was reduced, F_3 was perceived as F_2 .

Hedrick and Nabelek (2004) used synthetic full-spectrum vowels in a perception experiment in which the amplitude of F_2 was decreased in a nine-step continuum from /u/ to /i/. The results indicate that, for normal hearing listeners, a decrease in the amplitude of F_2 resulted in a greater number of /i/ responses. In studying vowel perception, the use of full-spectrum vowels could be problematic because when F_2 amplitude is low relative to the amplitude of the harmonics between formants, formant peaks simply lose contrast as the peak becomes less sharp. With full-spectrum stimuli, it may be difficult to distinguish if a change in vowel perception is due to a change in the spectral saliency of the peak, or due to higher frequency formants being masked by lower formants.

The purpose of this research is to investigate what role formant amplitude plays in the perception of vowels for speakers of English. Though studies have manipulated either F_2 or F_3 individually, this study varied the amplitude of both F_2 and F_3 in a fully-crossed design. This is done in order to separate out the perceptual effect of spectral tilt, which is a potential confound in synthetic monophthongs (Kieffe & Kluender, 2005).

Because changes in identification could be due to loss of spectral sharpness or simultaneous masking, this study used two sets of stimuli to resolve this problem. Full-spectrum stimuli were synthesized normally while incomplete-spectrum stimuli consisted of harmonics at the centre of the first five formants, as well as one harmonic on either side.

2. METHOD

Figures 1 (a) and (b) show spectra of the /u/ and /i/ endpoints of the full-spectrum continuum, respectively. Likewise, figures 1 (c) and (d) show spectra of the /u/ and /i/ endpoints of the incomplete-spectrum continuum.

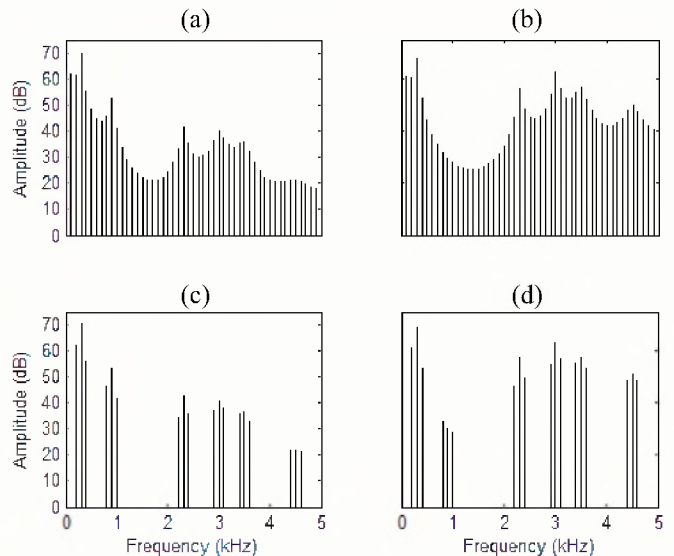


Fig. 1. Spectra of stimuli used in Experiment I.

Eligible participants completed a forced-choice vowel identification task with two sets of stimuli: full-spectrum and incomplete-spectrum. Participants labeled stimuli as either /i/ or /u/. The stimuli varied in amplitude of F_2 and F_3 in a fully-crossed 7x7 design. Each participant identified each stimulus 8 times. Each stimulus in the 7x7 design was pre-

sented randomly following the phrase “you will now hear the vowel...”.

3. RESULTS

The following figures present the number of /i/ responses of 14 participants for the full-spectrum and incomplete-spectrum, respectively. The maximum number of possible responses for each stimulus is 112.

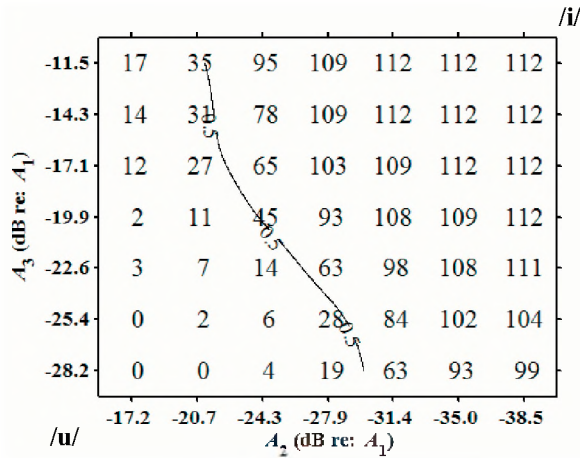


Fig. 2. Number of /i/ responses for the full-spectrum stimuli.

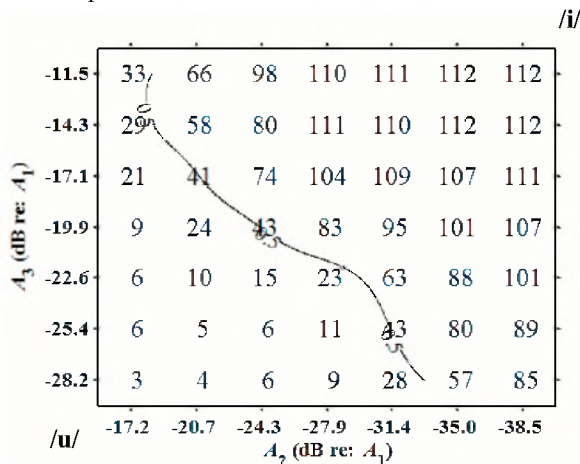


Fig. 3. Number of /i/ responses for the full-spectrum stimuli.

Overall, there were a greater number of /i/ responses in the full-spectrum condition. It is possible that this is a result of reduced spectral contrast. A potential explanation for this is that reduced spectral contrast caused listeners to perceive F_3 as F_2 , and identify the stimulus as /i/.

There were more /u/ responses in the incomplete-spectrum condition. One possibility is that listeners were able to perceive F_2 even at very low amplitudes.

Logistic regression was performed on the responses from each subject (McCullagh & Nelder, 1989). Regression coefficients were then subjected to within-subjects t -test (Davis, 2002). There was a significant difference in coeffi-

cients for F_2 amplitude between the two conditions ($p = .001$). However, there was no significant difference in coefficients for spectral tilt between the two conditions.

4. DISCUSSION

Results from this experiment suggest that several perceptual mechanisms may account for shifts in vowel identification with changes in formant amplitude. First, formant amplitude contributes to the global spectral shape of the vowel, which has been shown to influence vowel perception in synthetic monophthongs (e.g., Kiefe and Klunder, 2005). However, differences in response patterns between the two conditions suggest that spectral contrast in the region of F_2 also contributes to vowel identification. It is likely that simultaneous masking plays a role as well.

Future studies will continue to examine the role of F_2 in vowel perception using a signal detection task. Participants will be presented with stimuli that are identical in formant frequency and amplitude, except that in some pairings the second formant has been removed. Three different pairings will be used: identical, unaltered stimuli; one unaltered stimulus followed by one altered stimulus; one altered stimulus followed by one unaltered stimulus. The stimuli will be separated by 250 ms of silence. Participants will be asked to choose if the stimuli are the same or different.

REFERENCES

- Aaltonen, O. (1985). The effect of relative amplitude levels of F_2 and F_3 on the categorization of synthetic vowels. *Journal of Phonetics*, 13, 1-9.
- Ainsworth, W.A. & Millar, J.B. (1972). The effect of relative formant amplitude on the perceived identity of synthetic vowels. *Language and Speech*, 15, 328-341.
- Davis, C.S. (2002). *Statistical Methods for the Analysis of Repeated Measurements*. New York: Springer.
- Hedrick, M.S. & Nabelek, A.K. (2004). Effect of F_2 intensity on identity of /u/ in degraded listening conditions. *Journal of Speech, Language, and Hearing Research*, 47, 1012-1021.
- Ito, M., Tsuchida, J. & Yano, M. (2001). On the effectiveness of whole spectral shape for vowel perception. *Journal of the Acoustical Society of America*, 110(2), 1141-1149.
- Kiefe, M. & Klunder, K.R. (2005). The relative importance of spectral tilt in monophthongs and diphthongs. *Journal of the Acoustical Society of America*, 117(3) 1395-1404.
- McCullagh, P. & Nelder, J.A. (1989). *Generalized Linear Models*, 2nd ed. London: Chapman & Hall.
- Miller, J.D. (1984). Auditory processing of the acoustic patterns of speech. *Archives of Otolaryngology*, 110, 154-159.
- Nabelek, A.K., Czyzewski, Z., & Krishnan, L.A. (1992). The influence of talker differences on vowel identification by normal hearing and hearing-impaired listeners. *Journal of the Acoustical Society of America*, 92(3), 1228-1246.
- Rosner, B.S. & Pickering, J.B. (1994). *Vowel Perception and Production*. Oxford: Oxford University Press.