Static and Dynamic Cues in Vowel Production in Hijazi Arabic

Wael Almurashi, Ghada Khattab and Jalal Al-Tamimi
(Newcastle University)
w.a.o.almurashi2@newcastle.ac.uk / ghada.khattab@ncl.ac.uk / jalal.al-tamimi@ncl.ac.uk

Static cues for vowel identification have been widely used, with vowel formant values being taken at midpoint (50%) in L1 acoustics research, as it is believed that they represent the best acoustic characteristics of monophthong vowels (Peterson & Barney 1952). Nevertheless, many studies (e.g., Nearey & Assmann 1986, among others) have reported that dynamic cues—in particular, vowel-inherent spectral change (VISC) and the three-point model (e.g., Hillenbrand et al. 1995, among others)—can obtain better identification of vowels by using discriminant analysis with the formant pattern sampled at multiple locations. To date, many studies on Arab L1 vowel production have concentrated on the static acoustic features (e.g., Almbark 2012), and the role of dynamic cues has not been fully investigated (although see Al-Tamimi 2007). Therefore, this research presents an acoustic description of the vowel production of L1 Hijazi Arabic (HA) to evaluate the importance of static and dynamic cues, particularly VISC models (direction, slope, and offset) and three sample-points model, in the classification of HA vowels. An added purpose is to estimate the extent to which vowel duration and F3 aid in the classification accuracy.

Recordings were made of 12 male HA speakers producing eight HA vowels in /hVd/ syllables that were included in a carrier sentence. The first three formant frequencies were measured during the vowel duration three times—once (at 50%) for the static model; twice (at 20% and 80%) for the VISC models; and three times (at 20%, 50% and 80%) for three-point model —and then tested using discriminant analysis. The discriminant analysis results show that, in general, the three-point model had higher correct classification rates for all 8 HA vowels than the static model and VISC models. However, all models performed well in identifying HA lax and tense vowels and vowel pairs, particularly when the vowel duration was added. The average correct classification rate was 98% for the three-point model, 96% for the static model, and (86% for direction, 85% for slope, and 74% for offset) for the VISC models. The vowel duration was found to play a significant role in the classification accuracy (+11%), while the role of F3 appeared to have a minor influence on the classification's accuracy (+1%). In conclusion, the three-sample-point method was the most accurate model for classifying HA vowels in all stages, and the vowel duration was the most important cue on the classification accuracy of HA vowels.

References:

Almbark, R. (2012). The perception and production of SSBE vowels by Syrian Arabic learners: The foreign language model (Unpublished PhD Thesis). University of York, UK.

Al-Tamimi, J. (2007). Static and Dynamic cues in Vowel Production: a cross dialectal study in Jordanian and Moroccan Arabic. In proceedings of the *The 16th International Congress of Phonetic Sciences (ICPHS)*, 541-544. Saarbrücken, Germany.

Hillenbrand, J. M. Getty, L. A. Clark, M. J. & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *The Journal of the Acoustical Society of America*. 97: 3099-3111.

Nearey, T. M., and Assmann, P. F., (1986). Modeling the role of inherent spectral change in vowel identification. *The Journal of the Acoustical Society of America*, *80*(5), 1297-1308.

Peterson, G. E. & Barney, H. L. (1952). Control methods used in a study of the vowels. *The Journal of the acoustical society of America, 24*(2), 175-184.