Ability of different machine learning algorithms to predict the classification of second language vowels

Georgios P. Georgiou (University of Nicosia)

Machine learning techniques have been increasingly utilized over the past decade to predict second language (L2) speech perception patterns. This is based on the either direct or indirect assumptions of several speech models (e.g., Speech Learning Model/SLMr; Flege, 1995, Flege & Bohn, 2021; Perceptual Assimilation Model/PAM-L2, Best, 1995; Best & Tyler, 2007; Second Language Linguistic Perception model; Escudero, 2009; Universal Perceptual Model, Georgiou, 2021) that acoustic/articulatory-phonetic similarity between first language (L1) and L2 sounds can predict the perception of L2 sounds. However, predictions have largely depended on the use of discriminant analysis approaches (e.g., Escudero et al., 2012; Georgiou, 2022). To date, no studies have compared the classification accuracy of different algorithms. This study aims to assess the ability of three machine learning algorithms, namely, linear discriminant analysis (LDA), decision tree (C5.0), and artificial neural network (NNET) to predict the classification of L2 sounds in terms of L1 categories. The models were trained using the first three formants and duration of the L1 Cypriot Greek vowels /i e a o u/ and fed with the same acoustic features of the L2 Standard Southern British English (henceforth English) vowels /ɪ iː e ɜː æ ɑː ʌ ɒ ɔː uː ʊ/. The trained data were produced by 22 (*n*females =11) adult Cypriot Greek speakers. The vowels were embedded in a /pVs/ context and were part of the carrier phrase 'Léne <target word> tóra'. The testing data included the same acoustic parameters for English vowels and were produced by 20 (*n*females = 10) adult English speakers. The vowels were included in an /hVd/ context and as part of the carrier phrase "They say <word> now". The output was normalized using the Lobanov method. To validate the algorithms' accuracy, adult Cypriot Greek speakers of L2 English completed a perceptual classification task. The participants were 20 Cypriot Greek speakers (*n*females = 10) with an age range of 19-43 (*M*age = 31.9, *SD* = 6.93). They were born, raised, and permanently resided in Cyprus at the time of the study and had never lived in an English-speaking country for a long time. The results indicated that NNET predicted with success the classification of all L2 vowels with the highest probability in terms of L1 categories, while LDA and C5.0 missed only one vowel. NNET was also the most accurate in predicting the full range of above chance responses followed by LDA and C5.0. Overall, NNET demonstrated the best discrimination accuracy slightly followed by LDA, while C5.0 did not meet expectations. The models' plots are illustrated in Figures 1-3. The findings can have significant implications for updating the current speech acquisition theories by better understanding the role of acoustic cues in L2 speech perception and for better modelling of listeners' L2 sound discrimination difficulties using a classifier that offers optimal predictions.

Keywords: machine learning algorithm, sound classification, speech perception,

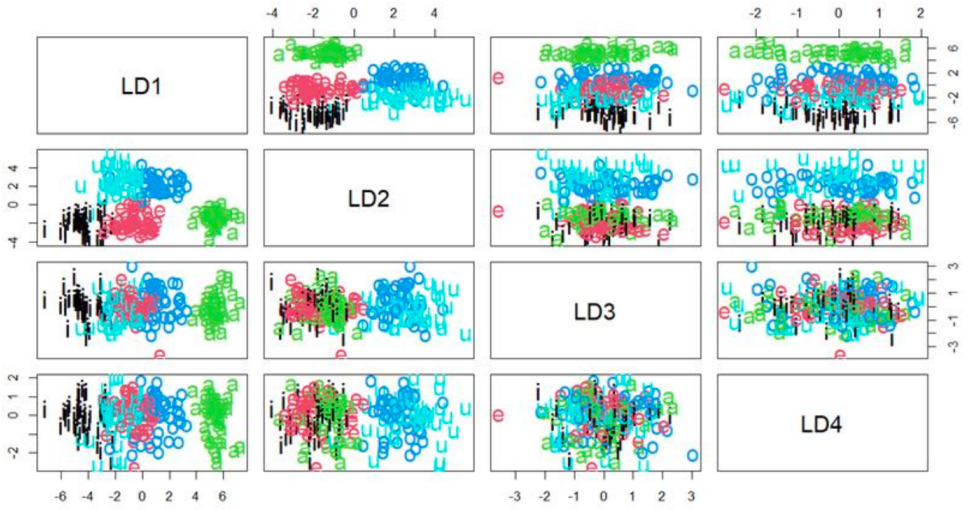

Figure 1: Linear Discriminant Analysis scatter plot for vowel classification.

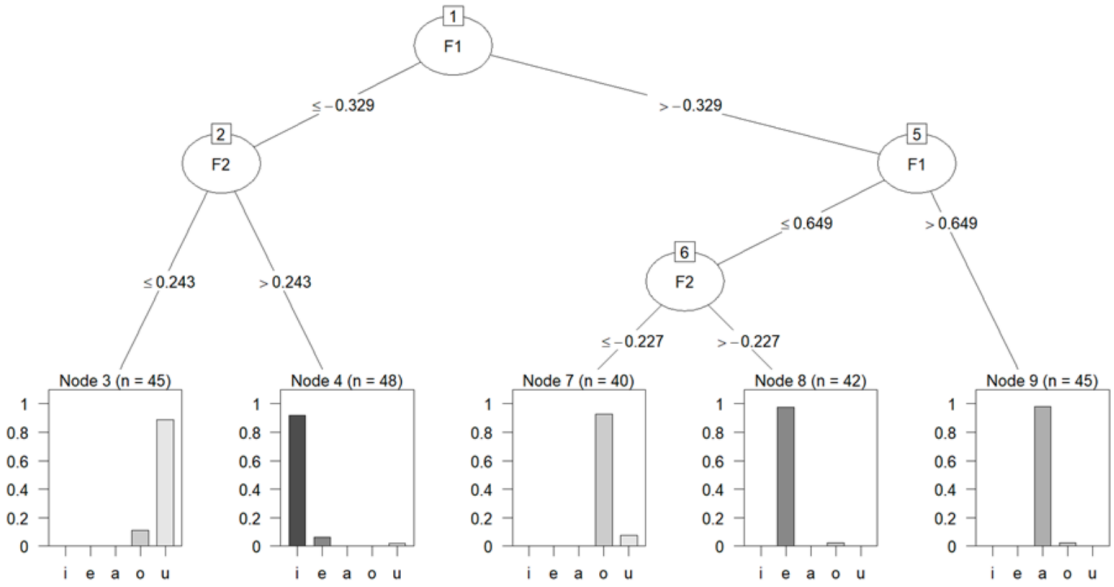

Figure 2: Decision tree for the C5.0 algorithm.

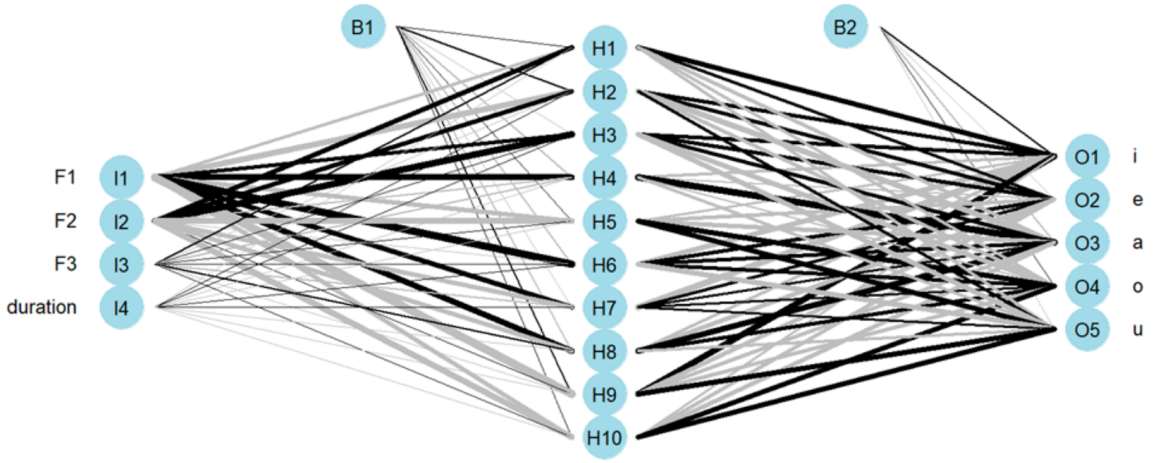

Figure 3: NNET architecture plot for the trained vowels.

References

Best, C. T. (1995). A direct realist view of cross-language speech perception: New Directions in Research and Theory. In W. Strange (Ed.), *Speech perception and linguistic experience: Theoretical and methodological issues* (pp. 171-204). Baltimore: York Press.

Best, C. T., & Tyler, M. (2007). Non-native and second-language speech perception: Commonalities and complementarities. In O-S. Bohn & M. J. Munro (Eds.), *Second language speech learning: In honor of James Emil Flege* (pp. 13-34). Amsterdam/Philadelphia: John Benjamins.

Escudero, P. (2009). Linguistic perception of "similar" L2 sounds. In P. Boersma and S. Hamann (eds), *Phonology in Perception* (pp. 151–190). Mouton de Gruyter.

Escudero, P., Simon, E., & Mitterer, H. (2012). The perception of English front vowels by North Holland and Flemish listeners: Acoustic similarity predicts and explains cross-linguistic and L2 perception. *Journal of Phonetics, 40*(2), 280-288.

Flege, J. E. (1995). Second language speech learning: Theory, findings and problems. In Strange, W. (Ed.), *Speech perception and linguistic experience: Theoretical and methodological issues* (pp. 233-277). Baltimore: York Press.

Flege, J. E., & Bohn, O. S. (2021). The revised speech learning model (SLM-r). *Second language speech learning: Theoretical and empirical progress*, 3-83.

Georgiou, G. P. (2021). Toward a new model for speech perception: The Universal Perceptual Model (UPM) of Second Language. *Cognitive Processing, 22*(2), 277-289.

Georgiou, G. P. (2022). The acquisition of /ɪ/–/iː/ is challenging: Perceptual and production evidence from Cypriot Greek speakers of English. *Behavioral Sciences, 12*(12), 469.