How to Start Hearing Things, or The Role of Coarticulation and Fine Phonetic Detail in Speech Perception

Corresponding Author (uliana.sudakova@etu.u-paris.fr)[1], Uliana Sudakova[1], Jalal Al-Tamimi[1]

[1]Université Paris Cité, CNRS, LLF, Paris, France

Introduction: Speakers constantly monitor their speech and choose particular wordforms, intonation, and phonetic forms depending on the addressee, communicative goals, cognitive load etc. [1, 2]. Following the "Adaptive Variability" and "Hypo-/Hyper-speech" theories [4, 5], speakers produce only as much phonetic precision as the situation demands. Hypo-speech is maximally reduced and economical, while hyper-speech is maximally expanded, leading to great articulatory precision. Alongside this continuum, vowel identification can pose challenges. Vowel sustained area was assumed to hold crucial information for its identification [8]; however, in CVC syllables the "midsyllable formant values of coarticulated vowels often do not reach the static steady-state values of vowels produced in isolation" [8], leading to clear confusability. Using a silent-center paradigm, in which the center portion of the vowel is deleted, the authors showed that participants could still identify the vowel successfully, highlighting an important role for vowel onset and offset in vowel identification. We aim to examine whether vowel identification will be possible if the sustained area of the vowel is deleted from the CV sequence and when changing from a hypo to hyper-speech mode.

Method: British English monophthongs preceded by alveolar and velar stops in /həCV/ sequences will be recorded, in both hypoarticulated form in the carrier phrase "I say the word /həCV/ some more" and in the hyperarticulated form via citation-form list reading [3, 6, 7]. The selected stimuli (point vowels /i, ɑ, u/) will then be used in a perception experiment (a forced-choice identification task) testing whether participants can be trained to identify the following vowel from the consonantal burst alone. Participants will be randomly assigned to three groups: hyper-to-hypo training (G1), hypo-to-hyper training (G2), or a no-training control (G3). The experiment has three blocks: a pre-training block establishes baseline accuracy (all groups), a training block (G1 and 2) presents stimuli along a progression of decreasing acoustic information: full CV sequence; burst + transition; burst only, with G1 moving from hyperarticulated to hypoarticulated tokens and G2 in the reverse direction; trial-by-trial feedback is given. A post-training block, identical but without feedback, measures learning transfer. Stimuli will be presented in random order with three repetitions. Accurate category responses will be analysed using logistic mixed-effects regression.

Expected results: We predict that training will improve vowel recognition from consonantal cues, with differences in trainability across vowels; decreased accuracy as acoustic information is reduced; an advantage for hyperarticulated sequences; and that CV sequences in velar context will be better identified than alveolars. Results will be presented at the conference.

References

[1] Denes, P. B., Pinson, E. N. The speech chain: The physics and biology of spoken language. Anchor Press. 1973.

[2] Farnetani, E., Recasens, D. Coarticulation and connected speech processes. In W. J. Hardcastle, J. Laver, & F. E. Gibbon (Eds.), The handbook of phonetic sciences (2nd ed., pp. 316–352). Wiley-Blackwell. 2010.

[3] Jenkins, J. J., Strange, W., Miranda, S. Vowel identification in mixed-speaker silent-center syllables. Journal of the Acoustical Society of America, 1994. 95(2). 1030–1043. https://doi.org/10.1121/1.410014

[4] Lindblom, B. Economy of speech gestures. In P. F. MacNeilage (Ed.), The production of speech. Springer, 1983. 217–245.

[5] Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H&H theory. In W. J. Hardcastle & A. Marchal (Eds.), Speech production and speech modelling. Kluwer Academic,1990. 403–439.

[6] Stack, J. W., Strange, W., Jenkins, J. J., Clarke, W. D., Trent, S. A. Perceptual invariance of coarticulated vowels over variations in speaking rate. Journal of the Acoustical Society of America, 2006. 119(4). 2394–2405. https://doi.org/10.1121/1.2171837

[7] Strange, W. Dynamic specification of coarticulated vowels spoken in sentence context. Journal of the Acoustical Society of America, 1989. 85(5). 2135–2153. https://doi.org/10.1121/1.397863

[8] Strange, W., Jenkins, J. J. Dynamic specification of coarticulated vowels: Research chronology, theory, and hypotheses. In G. S. Morrison & P. F. Assmann (Eds.), Vowel inherent spectral change. Springer. 2013. 87–115.