Vowel production in congenitally blind and sighted Australian English speakers

Milica Janić[1], Teja Rebernik[1], Pauline Veenstra[1], Emma Wissink[1], Martijn Wieling[1,2], Michael Proctor[3]
[1]University of Groningen, [2]Haskins Laboratories, [3]Macquarie University

The influence of visual deprivation on speech production has not been studied extensively. Previous research (e.g., Ménard, Dupont, Baum & Aubin, 2009; Ménard et al., 2013), which focused on French, found that sighted speakers produce vowels that are spaced further apart in the acoustic space than blind speakers. However, these results have not yet been replicated in another language, and one study indicated the opposite pattern for Dutch blind speakers (Veenstra, Everhardt & Wieling, 2018).

The goal of our study was to investigate how the absence of visual feedback impacts vowel production in congenitally blind and sighted Australian English speakers. In line with the findings of Ménard, Dupont, Baum & Aubin (2009), we hypothesise that vowel production in sighted speakers will be spaced further apart than in congenitally blind speakers. Specifically, this would result in blind speakers having a smaller vowel space area than sighted speakers.

Method and participants

We recorded acoustic data from 10 congenitally blind (4 male) and 10 sighted (4 male) native speakers of Standard Australian English (AusE), who matched as closely as possible for age (ages ranged between 30-65 for blind speakers and 27-65 for sighted speakers). The congenitally blind participants had different etiologies of blindness but were all blind since birth and never able to see more than shapes or light. 12 AusE monophthongs and 6 diphthongs were elicited in three stressed contexts: word-initially (*V-word*), in a *hVd* word, and in isolation (e.g., "***E****ven* as in *h****ee****d* as in ***ee***" for vowel /i:/). Stimulus sentences were presented twice (on a laptop screen for sighted and Braille cue cards for blind speakers) in random order, to elicit 6 repetitions of each vowel per speaker. A subset of the vowels is analyzed here (see below).

Analysis

Speech recordings were annotated in PRAAT (Boersma & Weenink, 2019), and the acoustic limits of target vowels for each speaker were manually located in order to ensure adequate formant tracking. The first four formants were automatically tracked over each vocalic interval, and formant estimates F1 to F3 were extracted at midpoint of vowel duration and mel-transformed. We calculated vowel space areas (VSA) with the first and second formants (in mels) using the *vowelMeansPolygonArea* function in the *phonR* package (McCloy, 2016). The vertices of the vowel space were defined by vowels /i:/, /æ/, /ɐ:/ and /o:/ (see Figure 1). A VSA was calculated for each vowel context per participant.

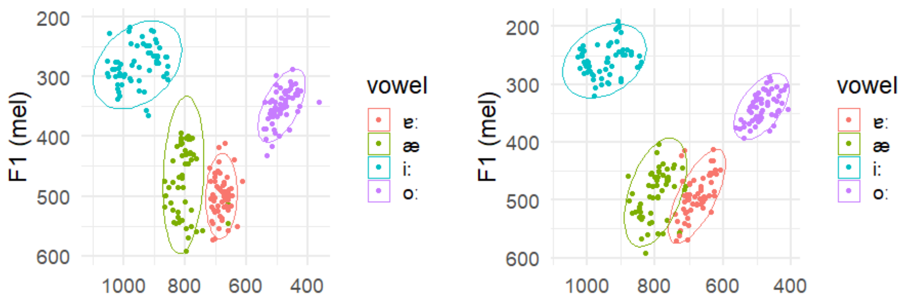

Figure 1: Acoustic vowel spaces for blind (left) and sighted (right) speakers of Australian English. Ellipses (outline at 2 SD) centred on mean formant values for each vowel.

Results

To assess the potential effect of group (BLIND versus SIGHTED) on the size of the vowel space area, we built linear mixed-effects models, with group and gender as fixed effects and the speaker as a random effect. Due to the limited sample size and composition, we were not looking for significant *p*-values but rather focusing on effect sizes (only medium to large effect sizes are highlighted here, i.e. Cohen's $d \geq 0.5$).

We observed a medium effect of group on VSA (d = 0.5), indicating that the congenitally blind speakers of Australian English have a smaller VSA than the sighted speakers. As expected, there was a large effect of gender (d = 1.7; larger VSA for women), however there was no meaningful interaction between gender and group. Adding age as a fixed effect did not improve the model (there was only a small effect of age on VSA; d = 0.2). Figure 2 depicts the mean VSA values per group and gender.

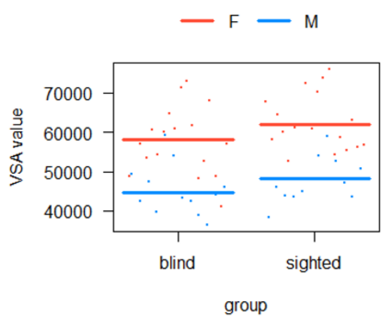

Figure 2: VSA (in mel^2) per group and gender.

Discussion

Although preliminary, these results are consistent with our hypothesis, and with findings of previous studies of congenitally blind speakers: the Australian English blind speakers in this study produce vowels that are clustered closer together than those produced by sighted speakers. This may indicate that a lifelong absence of visual feedback has an impact on vowel production. However, while the pattern is similar, these data cannot be directly compared with earlier studies of Canadian French because of differences in the properties of the vowels and the organization of the vowel spaces. In particular, there are important differences in the relationships between rounding and backness between French and English, for example. Further investigations into the distribution of F3, and F3-F2 differences for key vowel pairs will shed more light on these issues.

References

Boersma, P., & Weenink, D. (2019). Praat: doing phonetics by computer [Computer program]. Version 6.1.08, retrieved 5 December 2019 from http://www.praat.org

McCloy, D. R. (2016). *phonR: tools for phoneticians and phonologists.* R package v. 1.0-7.

Ménard, L., Dupont, S., Baum, S. R., & Aubin, J. (2009). Production and perception of French vowels by congenitally blind adults and sighted adults. *JASA*, *126*(3), 1406-1414.

Ménard, L., Toupin, C., Baum, S. R., Drouin, S., Aubin, J., & Tiede, M. (2013). Acoustic and articulatory analysis of French vowels produced by congenitally blind adults and sighted adults. *JASA*, *134*(4), 2975-2987.

Veenstra, P., Everhardt, M. K., & Wieling, M. (2018). *Vision deprived language acquisition: Vowel production and ASR efficacy*. Poster presented at the 16th Conference on Laboratory phonology (LabPhon), Lisbon, Portugal.