How revealing are vowels for the sex of children?
A perception and production study

The purpose of this study was to test to which extent the sex of 6-7-year-old children can be distinguished acoustically and perceptually from their vowels. Previous research has shown that listeners usually recognize the sex of children above chance level despite f0 not being a discriminating cue [1], with increasing accuracy for older children [1,2,3] and stimuli of greater duration [4]. Formant distinctions between male and female vowels have also been shown to be more marked in open and high front than in high back vowels [5,6,7,8,9, etc.]. A further aim was thus to verify whether sex identification was vowel-dependent.

The speech material came from 37 Albanian-speaking children (18F). Sex-specific acoustic differences in the vowel space were observed, which are similar to those of adults (Fig1). For the perception experiment, the stressed vowel was extracted from five words containing either /i,u,e,o,a/. The listeners, 111 German-speaking adults, were randomly assigned to one of five conditions, each containing tokens of one vowel only. Listeners responded on a two-forced-choice scale (*Mädchen* 'girl', *Junge* 'boy').

Overall, the correct identification rate was 65%, above chance level (z=4.8, p<.001) and similar to [1,4]. We found a significant interaction [10,11,12] between vowel and sex (X^2[10]=111.4, p<.001). Post-hoc tests (Fig2) showed that males and females were equally well recognized in /i,e,o/, but there was a bias towards identifying males in /u,a/ (z=3.7, p=.007; z=5, p<.001). For boys, the predicted correct identification rate was particularly high for /a/ (86%), and particularly low for girls for /u/ (47%) and /a/ (54%).

German-speaking adult listeners were globally successful at recognizing the sex of Albanian-speaking children when listening to vowels only. The misidentification of females for /u/ and /a/ could be explained by acoustic similarity in formants and low vowel-intrinsic f0, respectively. Future work will expand on cross-linguistic and stimulus length issues.

(300 words)

References

[1] Perry, T. L., Ohde, R. N., & Ashmead, D. H. (2001). The acoustic bases for gender identification from children's voices. *Journal of the Acoustical Society of America*, *109*(6), 2988–2998.

[2] Gelfer, M. P., & Bennett, Q. E. (2013). Speaking fundamental frequency and vowel formant frequencies: Effects on perception of gender. *Journal of Voice*, *27*(5), 556–566.

[3] Ingrisano, D., Weismer, G., & Schuckers, G. H. (1980). Sex identification of preschool children's voices. *Folia Phoniatrica*, *32*(1), 61–39.

[4] Günzburger, D., Bresser, A., & ter Keurs, M. (1987). Voice identification of prepubertal boys and girls by normally sighted and visually handicapped subjects. *Language and Speech*, *30*(1), 47–58.

[5] Peterson, G. E., & Barney, H. L. (1952). Control methods used in a study of the vowels. *Journal of the Acoustical Society of America*, *24*(2), 175–184.

[6] Simpson, A. P. (2009). Phonetic differences between male and female speech. *Language and Linguistics Compass*, *3*(2), 621–640.

[7] Pettinato, M., Tuomainen, O., Granlund, S., & Hazan, V. (2016). Vowel space area in later childhood and adolescence: Effects of age, sex and ease of communication. *Journal of Phonetics*, *54*, 1–14.
[8] Vorperian, H. K., & Kent, R. D. (2007). Vowel acoustic space development in children: A synthesis of acoustic and anatomic data. *Journal of Speech, Language, and Hearing Research*, *50*(6), 1510–1545.
[9] Diehl, R., Lindblom, B., Hoemeke, K. & Fahey, R. (1996). On explaining certain male-female differences in the phonetic realization of vowel categories. *Journal of Phonetics*, *24*(2), 187-208.
[10] R Core Team. (2021). *R: A language and environment for statistical computing*. https://www.R-project.org
[11] Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48.
[12] Lenth, R., Love, J., & Herve, M. (2018). *emmeans: Estimated marginal means, aka least-squares means* (R package version 1.1.2).

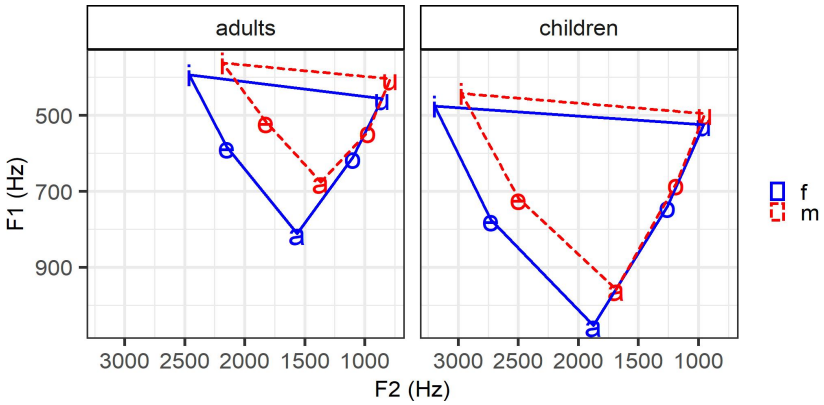

Fig1. Five oral vowels of Albanian in F1/F2 planes averaged across 28 control adults and 37 children for male (red dashed) and female (blue solid) speakers. A subset of the child data was used as stimuli in the perception experiment

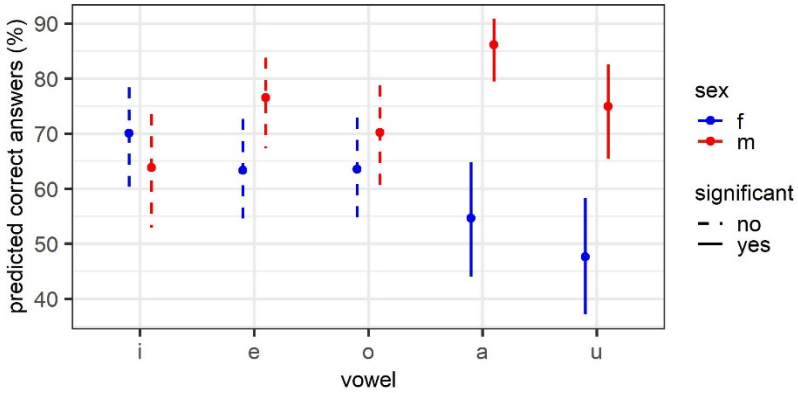

Fig2. Predicted rate of correct answers per vowel and sex, where dashed and solid lines represent non-significant and significant differences between sexes respectively.