Gender-based studies on the spoken language of children

Rosa Franzke,[1] Katrin Wolfswinkler[1], Jonathan Harrington[1]

[1]Institut für Phonetik und Sprachverarbeitung, LMU München

rosa.franzke@campus.lmu.de, katrin@phonetik.uni-muenchen.de, jmh@phonetik.uni-muenchen.de

The purpose of this study was to combine perception, acoustics and anatomy of children´s speech productions in order to assess the characteristics that account for gender differences in 6-7-year-olds.

Generally, the differences between male and female speech that lead to sex differences are in large part due to different anatomical structures such as vocal fold size and vocal tract length. That is, an average adult male has a longer vocal tract and hence lower formant frequencies than an average adult female (Ladefoged and Broadbent, 1957; Laver and Trudgill, 1979; Peterson and Barney, 1952). As for children though, the specific age at which these distinctions start to occur is controversial (Lee et al., 1999).

There have been several investigations that found evidence for lower vowel resonances for boys than for girls, e.g. Bennett (1981) and Perry et al. (2001) for American children, and Busby and Plant (1995) for Australian children. However, the reasons for these acoustic differences between the sexes in early childhood remain largely unexplored. This research seeks to resolve this issue by exploring not just acoustic but also the perceptual and physiological bases of sex differences in the speech of German children. For this purpose we considered the following issues:

1) Perception: Are adult listeners able to perceive correctly the gender of German speaking children as young as 6-7 years of age?
2) Acoustics: Do fundamental and vowel formant frequencies differentiate gender in 6-7-year-old girls and boys and what is their role in the perception of sex?
3) Anatomy: Are there already anatomical differences between these girls and boys that could cause potential differences in 1) and 2)?

The acoustic data was taken from a database of child speech consisting of productions by 25 first graders aged between 6-7 years in two primary schools in the surroundings of Altötting, located in a West-Central-Bavarian dialect area 90 km from Munich. For the database, 58 mostly trochaic words were pronounced 4 times by the 25 children in a picture-naming task. For the current study, 8 different disyllabic words with front and back vowels (i.e. /i/, /e/, /o/, /u/) in stressed position from 10 different children (5 f, 5 m) were selected from this database.

The first task was to assess whether listeners could reliably distinguish between sex in children. For this purpose, an online perception experiment was conducted in which 94 adult listeners provided a forced choice response (male or female) to each of the 8 productions of the 10 children, which were repeated once (160 stimuli in total per listener). The listeners, who were recruited via a mailing list, were all registered students of the LMU in Munich. The results showed that these adult listeners were able to identify the sex of 6-7-year-old children with an identification rate of 71.6% which is significantly ($p < .01$) greater than chance.

We then sought to establish the acoustic basis of the perceptual responses in two separate investigations. Firstly, a correlation of the perceptual responses with formants and f0 of the perceived utterances. Secondly, an analysis of the same production data in order to examine whether girls and boys differed on the same acoustic cues as the listeners relied upon for their decision. The results of both analyses broadly converged in showing that f0 was ineffective and F3 the most important parameter for the sex distinction. More specifically, male children's productions were significantly ($p < .01$) associated with a lower F3 (Fig. 1). Compatibly (Fig. 2), listeners judgments were significantly influenced by F3 ($p < .01$) in the same direction i.e. they associated lower F3 with male productions.

We subsequently tested whether physiological measures between girls and boys could explain these differences in F3. For this purpose, measurements of head height, head width and distance between ear entrance and upper lip were extracted from video recordings of the children,

which were obtained in the course of ultrasound recordings. Tongue length as well as distance between the ultrasound sensor and the highest point of the tongue were extracted from the ultrasound data. Overall, there were no significant sex differences in any of the anatomical measurements (and none that correlated with sex-based F3 differences).
This leads us to conclude tentatively that the observed acoustic differences and associated perceptual responses came about for sociophonetic reasons rather than as a result of physical differences. This is consistent with Johnson´s (2005) finding that sex-based formant differences in adults vary from language to language, suggesting that such variation cannot be purely anatomical. Our present study suggests that learning to sound male or female can be observed in children's first language already at the age of 6-7 years.

Keywords: perceptive gender recognition, gender-specific acoustics, gender-specific anatomy, voice and language of prepubescent children, correlation between perception, acoustics & anatomy

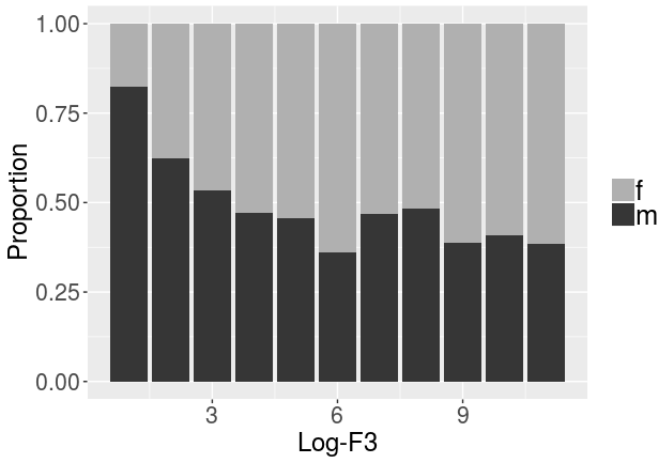

FIG. 1. Logarithmized F3 grouped into 11 bins with increasing height from left to right (x-axis) as a function of the proportions of female vs. male speakers (y-axis)

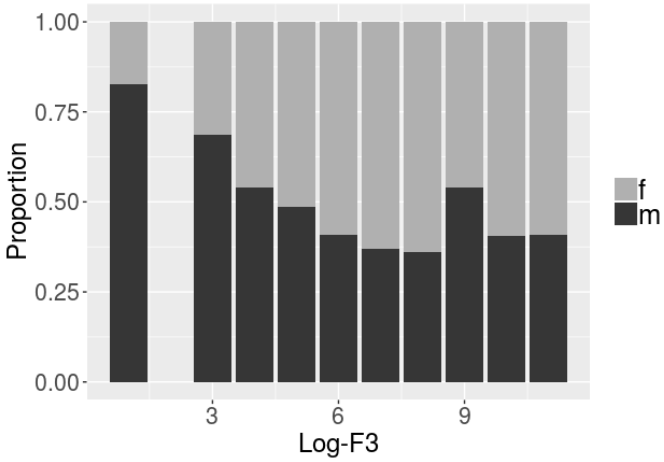

FIG. 2. Logarithmized F3 grouped into 11 bins with increasing height from left to right (x-axis) as a function of the proportions of female vs. male answers of the listeners (y-axis).

References

Bennett, S. (1981). Vowel formant frequency characteristics of preadolescent males and females. *The Journal of the Acoustical Society of America* 69, 231.

Busby, P. A. und Plant, G. L. (1995). Formant frequency values of vowels produced by preadolescent boys and girls. *The Journal of the Acoustical Society of America*, 97(4), 2603-6.

Johnson, K. (2005). Speaker normalization. In R. Remez, & D. B. Pisoni (Eds.), *The Handbook of Speech Perception*. Blackwell, 363 – 89.

Ladefoged, P. und Broadbent, D. E. (1957). Information conveyed by vowels. *The Journal of the Acoustical Society of America* 29, 98-104.

Laver, J., and Trudgill, P. (1979). 'Phonetic and linguistic markers in speech in K. Scherer and H. Giles (eds.) *Social Markers in Speech.* Cambridge University Press, Cambridge, 1–32.

Perry, T. L., Ohde, R. N und Ashmead, D. (2001). The acoustic bases for gender identification from children's voice. *The Journal of the Acoustical Society of America 109*(6), 2988-98.

Peterson, G. E. und Barney, H. L. (1952). Control methods used in a study of the vowels. *The Journal of the Acoustical Society of America*, 24, 175-84.