Effects of vocal tract growth on gender and vowel identification based on simulated children’s vowels

Brad H. Story¹, Houri K. Vorperian², & Kate Bunton¹

¹Department of Speech, Language, and Hearing Sciences, University of Arizona
²Waisman Center, University of Wisconsin-Madison

Introduction

The anatomical structures that form the vocal tract exhibit non-uniform growth during the course of development from infancy to adulthood (Vorperian et al., 2009). These growth patterns also indicate sex differences, even at prepubertal stages (Vorperian et al., 2011). It is not well understood how these non-uniform scalings of the vocal tract length affect the formant frequencies, and consequently how they affect both vowel and gender identification.

The purpose of this study was to use a model of the vocal tract to simulate vowels produced by a male talker and female talker at ages ranging from 1 to 20 years, and present them to listeners who were asked to identify both the vowel and gender of the talker.

Method

Speech Sample

Set 1: 160 simulated vowels
- Generated with TubeTalker (Story, 2013) using a nonlinear vocal tract length (VTL) growth algorithm
- American English vowels /i, æ, o, u/
- Male vs female growth patterns imposed on male and female vocal tract shapes, respectively
- Ages 1 to 20 years
- Female contour rising and falling with a mean of 280 Hz
- Vowel duration = 500 ms

Set 2: 160 simulated vowels
- Identical to Set 1 except the overall VTL was uniformly adjusted.

Listeners

16 naïve listeners (13 F, 3 M)
- Mean age = 24.5 years
- English as native language
- Passed a hearing screening (0.5, 1, 2, & 4 kHz @ 20 dB)
- No professional experience with children

Listening Tasks
- Forced choice vowel id task based on 12 American English vowels using Alvis 3 interface (Hillenbrand, Gayvert, & Clark, 2015)
- Identify which of two vowel samples was produced by a boy (paired male vs. female, age matched).

Model and Simulation

Development of the vocal tract model was based on measurements of sex-specific anatomic landmarks along the vocal tract centerline (Fig. 1a). The measurements define five regions that grow by different amounts over the course of development (Fig. 1b). Vocal tract configurations produced by adjusting the length vector of vocal tract area functions for male/female vowel /i/ (Story, 2005; 2008) are shown at ages 2, 6, 10, and 20 years in Fig. 1c.

Results

The results of the vowel identification experiment for both nonuniform and uniform VTL growth across the entire age range are shown in the two confusion matrices below (Tables 1 & 2). In both cases, the /i/ and /u/ targets were identified as these same vowels more than 70% of the time, whereas identification of the /æ/ and /o/ targets was more spread among neighboring vowels.

Discussion

Vowel identification accuracy relative to the targets was fairly high. The confusions of the /i/ and /u/ vowels with neighboring vowels is expected. In particular, the high rate of identification of the target /æ/ as /a/ is likely due to our listener pool from the Southwest US where there is little perceived difference between these vowels.

Even though listeners could determine vowel identity, they were not able to reliably identify the gender of the simulated vowels in either the nonuniform or uniform VTL growth condition.

There are a number of limitations of the study that likely contributed to the inability of listeners to identify gender:
- The vocal tract configuration was static. Articulatory timing may enhance gender specific characteristics but was not implemented in the simulations.
- Scaling of the cross-sectional areas along the VTL was not included in the vocal tract model.
- The voice source was identical for all male and female simulations, and for all ages. This was done deliberately to eliminate voicing cues.

Conclusions

Although the approach to modeling the growth of the vocal tract produced vowels in accord with published data and were easily identified, it did not, by itself, produce acoustic characteristics that were indicative of gender.

Acknowledgement

Research supported by NIH R01-DC011275, NIH R01-DC006282, NSF BCS-1145911

References


