



Modifying speech to children based on perceived developmental level: An acoustic study of adults' fricatives

Hannah Julien¹, Benjamin Munson¹, Jan Edwards², Mary E. Beckman³, & Jeffrey J. Holliday³

¹Department of Speech-Language-Hearing Sciences, University of Minnesota, Minneapolis, ²Department of Communicative Disorders and Sciences, University of Wisconsin, Madison, ³Department of Linguistics, Ohio State University, Columbus



Production-Perception Dynamics and Speech-Sound Development

- In order to acquire adult-like knowledge of the vowel and consonant sounds, children must learn how to parse the distributions of these sounds in different parametric phonetic spaces in terms of the sound systems of the languages being acquired (Beckman, Munson, & Edwards, 2007; Munson, Edwards, & Beckman, in press; Pierrehumbert, 2003)

- The experiment in this poster is part of a larger effort to modeling the emergence of low-level categories by looking at the dynamics of the relationship between caregivers' input to children and children's subsequent productions

- Phonemic contrasts are generally hyperarticulated in infant-directed speech (Burnham et al., 2002; Kuhl et al., 1997).

- Recently, Cristia (2009) showed that mothers' acoustic differentiation between /s/ from /ʃ/ was greater in speech to older (12- to 14-month-old) children than to younger (4- to 6-month-old) ones, and that the older children whose mothers had more-robust differentiation of the /s-/ʃ/ contrast in their production had better perception of the contrast than those that did not.

- Liu, Kuhl, & Tsao (2003) found that Mandarin-acquiring infants whose mothers produced a more-expanded vowel space perceived the /s-/ʃ/ better than those whose mothers did not

- Cristia and Liu et al.'s results suggest a much finer-grained relationship between production and perception in learner-input dyads. Perhaps adults perceive in fine detail the degree to which a child's production (mis-)matches the adult forms, and modify their speech to children accordingly as a form of corrective feedback.

- This work focuses specifically on fricatives, because...

- ...they are late-acquired than vowels
- ...they are often produced in error by children with speech-sound disorders
- ...they are the locus of interesting sociolinguistic variation

Our Research Questions

- Do adults produce speech differently in response to productions they judge to be inaccurate compared to ones they judge to be accurate?

- How continuous is the relationship between perceived accuracy and modification of the response?

- What predicts individual differences across talkers in their propensity to modify speech in response to perceived inaccurate productions?

References

Burnham, D., Kitamura, C., Vallier, C., & Ullmer, C. (2002). What's new? Pursuing 'On talking to babies and animals. *Science* 296, 1425.

Cristia, Alejandra (2009). Individual variation in infant speech processing: Implications for language acquisition theories. Unpublished doctoral dissertation, Purdue University.

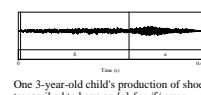
Kuhl, P.K., Andrusik, J.E., Chistovich, I.A., Chistovich, L.A., Kozhevnikova, E.V., Ryshina, V.I., Solyanova, E.I., Sundberg, U., Lacerda, F., 1997. Cross-language analysis of phonetic units in language addressed to infants. *Science* 277 (5326), 684-686.

Liu, H.-M., Kuhl, Patricia K., & Tsao, F.-M. (2003). An association between mothers' speech clarity and infants' speech discrimination skills. *Developmental Science*, 6, F1-F10.

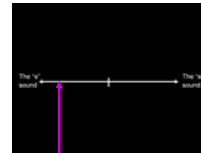
Urberg-Carlson, Kim, Kaiser, Eden, & Munson, Benjamin (2008). Assessment of children's speech production 2: Testing gradient measures of children's productions. Poster presented at the 2008 ASHA Convention, Chicago, 20-22. Downloaded on June 14, 2009 from http://www.xc.umc.edu/~munson08/Urberg-CarlsonEAL_Final.pdf.

The General Tactic

Listen to the initial CV of a child's attempt to say an /s/- or /ʃ/-initial word while looking at the picture the child was naming



Rate the child's production using a visual analog scale, (as in Urberg-Carlson et al., 2008)



Methods

Stimuli

- For the listen-rate-say (LRS) task, stimuli were 200 consonant-vowel sequences excised from the initial position of words produced by 2- and 3-year-old monolingual English-acquiring children in a picture-prompted repetition task. All of the target words had initial /s/ and /ʃ/ targets.

- Children's productions had been transcribed as being either correct, or being an [s]-for-ʃ/ or [ʃ]-for-[s] substitution.

Participants

- 22 adult native speakers of English with no specialized training in speech and language.

Procedures

- The LRS task is shown schematically above.

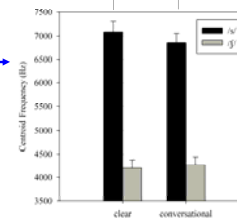
- The adults also completed a baseline task in which they produced the words from the LRS task embedded in sentences in conversational- and clear-speech styles

Analysis 1: Baseline Task

- The /s/ centroid frequencies were significantly higher in the clear-speech condition of the baseline task than in the conversational-speech condition. The /ʃ/ centroids were lower, though not statistically significantly so.

- Both /s/ and /ʃ/ were produced with longer durations in the clear-speech condition. The magnitude of the clear-speech effect was similar for /s/ and /ʃ/

- Vowels were also significantly longer and more-dispersed in the F1/F2 space in the clear-speech condition.



Acknowledgements

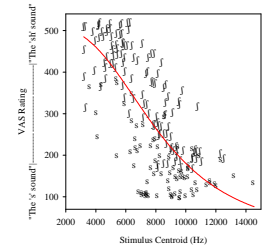
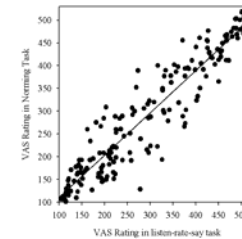
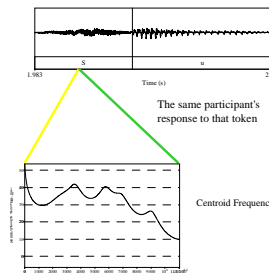
This research was supported by NSF Grant BC0872977 to Benjamin Munson, and by NIH grant R01 DC02912 to Jan Edwards. We thank Fangfang Li for conducting the acoustic analysis of the tokens. We thank Marie Meyer, Kari Urberg-Carlson, and Eden Kaiser for testing participants. We gratefully thank Gerald Butke and Marie Meyer for assistance with acoustic analyses. We thank Hans Heinz and Eric Foster-Lovner for help with this project.

Analysis 2: Accuracy Ratings and Stimulus Characteristics

- The VAS click locations were strongly correlated with those made by the listeners in Urberg-Carlson et al. (2008), despite the fact that the two tasks were different (below left)

- Listeners in this study knew the target, unlike the listeners in the previous study

- The VAS click locations were significantly correlated with the centroids of the fricatives in the stimuli, as in Urberg-Carlson et al. (2008). (Below right)



Analysis 3: Relating Perceived Accuracy to Adults' Productions

- The centroid frequency and duration of each of the fricatives from the LRS task was extracted. These were used as the dependent measures in four regressions: one predicting the /s/ centroids, one predicting the /ʃ/ centroids, one predicting the /s/ duration, and one predicting the /ʃ/ durations.

- The independent measures were /s/ or /ʃ/ centroid or duration for the same words in the conversational-speech baseline task, entered in the first step, and VAS rating for the word that was being named, entered on the second step.

- Centroid did not vary systematically as a function of the VAS rating that had been provided.

- For 7 of the 22 regressions, there was a significant relationship between VAS click location and duration of the /s/ and /ʃ/ responses. A representative talker's data is shown to the right.

- An additional 7 participants had a significant relationship for /ʃ/ but not for /s/.

- Hence, the majority of adults do modify their productions to children in response to productions they perceive to be inaccurate

