Child-Level Factors & Acquisition of the /t/-/k/ Contrast: Perception

Sara (Bernstein) Cline1
Allison Johnson2 Mary Beckman3 Jan Edwards2
and Benjamin Munson1

1Department of Speech-language Pathology, University of Minnesota, Minneapolis, Minnesota; 2Department of Communication Sciences and Disorders, University of Wisconsin, Madison; 3Department of Linguistics, The Ohio State University

Introduction

Background and motivation
- Children acquire the ability to produce speech sounds contrasts gradually (Cowan et al., 2001).
- Phonetic transcription in the traditional perceptual rating system
- Continuous rating scales such as those utilizing Visual Analog Rating (VAR) have been applied to the acquisition of phonetic contrasts (Galt & Leedy, 1984).
- VAR ratings correlate well with acoustic measures and are more gradient and hence potentially more informative than phonetic transcription (Cowan et al., 2001).

Aims of this study
1. Apply a continuous rating scale to characterize adult perceptions of children’s production for the /t/-/k/ contrast and derive measures of how reliably children’s productions differed based on these ratings.

Listener Participants
- Adult Listeners

Methods

Child Talkers
- Talker Participants
- 47 naive English speakers, aged 19-39 years
- No history of speech, language or hearing disorders
- Passed hearing screening, some late talkers, no other diagnoses

Hypothesized Predictor Variables
- Maternal Education Level
- Home Language Input
- Vocabulary

Stimulus Preparation
- Speech Recording
- Picture prompted auditory word repetition task
- A 3/t/-/k/ word list, 5/t/-/k/ initial words as part of longer word repetition task
- Recoded in a sound treated booth

Acoustic Event Tagging
- Initial consonant transcribed as:
  - [t]: more “t-like”
  - [k]: more “k-like”
  - [k]; intermediate more “b-like”
- Release of stop burst and onset of vocal fold vibration tagged in Praat
- Cross-sentential and cross-listener sequence extracted for 544 total tokens

Adult Listeners
- Listener Participants
- 47 native English speakers, aged 19-39 years
- No history of speech, language or hearing disorders
- Experiment split into 3 versions (about 20 tokens, 500 tokens each)
- Equal number of children represented in each version
- One talker presented across all versions
- Click only valid if no consistent vowel sound, sequence, clock location tagged automatically

Results: Listener Ratings

Category Differentiation
- Listener error differentiation showed that listeners rated all transcription categories differently (p<0.01) except for [t] vs. [k]
- Listeners rated transcriptions most widely dispensed along the VAR

Intra-rater Reliability
- Reliability was measured by the distance between clicks for replications/versions
- Reliability varied by experiment version, listener age, and listener sex
- Overall reliability was poor to good depending on listener; need for training?

Robustness of Contrast
- UtD Density plots of click locations for all productions showed a more robust /t/-/k/ contrast.
- Higher absolute slope = more robust contrast.
- The slope of each talker’s logit function was estimated.
- The slope of each talker’s logit function indicates robustness of contrast.
- Higher absolute slope = more robust /t/-/k/ contrast

Accuracy vs. Slope
- Accuracy = Present child’s attempts at a task that were transcribed as [t] or [k] (for each talker)
- Accuracy is a categorical measure and is not part of the regression analysis
- Accuracy provides an additional separation in data for tokens with high accuracy

Results: Listeners

Listener Description

- Listeners did not see a VAR to differentiate all transcription categories except [t] vs. [k] for /t/-/k/ contrasts.
- Reliability varied by experiment version and listener characteristics. Can reliable listeners be selected based on characteristics, or trained to be more reliable?

Set effects were present in the experiment versions; context of surrounding tokens influences listeners’ ratings.

The robustness of contrast measure “slope” characterizes the difference in VAR ratings for a /t/-/k/ contrast.
- Slope provides more information for tokens who have high protection accuracy.

Vocabulary size is significant in models predicting speech accuracy. Slope and Accuracy behave similarly in these models.

Conclusions

Acknowledgements

References

Mary E. Beckman, Benjamin Munson, and Jan Edwards.

1Department of Speech-language Pathology, University of Minnesota, Minneapolis, Minnesota; 2Department of Communication Sciences and Disorders, University of Wisconsin, Madison; 3Department of Linguistics, The Ohio State University

Acknowledgments

This research was supported by a grant from the National Institutes of Health (R01 DC050781 to Jan Edwards), a research grant from the University of Wisconsin-Madison’s Office of the Provost to Benjamin Munson, and an Educational Research Grant (Grant 09-51) provided by the American Speech-Language-Hearing Association to Mary Beckman. We also thank the children who participated in this study and their families for their time and energy in completing this research. We are appreciative of the support of the Children’s Learning Institute at Children’s Hospital of Philadelphia, where all data were collected, and to the Clinical Research Center at the University of Wisconsin-Madison for providing laboratory space and equipment support.