



Amy Andrzejewski, Eun Jong Kong and Jan R. Edwards  
University of Wisconsin-Madison

## INTRODUCTION AND RATIONALE

- Traditional studies of children's speech sound acquisition have relied primarily on phonetic transcription.
- However, listeners can be biased by various speaker-related factors.
- This biasing can go in either direction:
  - Listeners may be more lenient when they think they are listening to a younger child (e.g. Schellinger, 2008).
  - Listeners may be more critical when they think they are listening to a child with a speech sound disorder (e.g. Podol & Salvia, 1976).
- Much of this research has focused on biasing listeners with pictures of children with and without disorders.
- The focus of this research was to determine whether we could bias listeners with only an auditory prompt, as this is more ecologically valid.

## PREVIOUS RESEARCH WITH AUDITORY PROMPTS

- Schellinger and colleagues also used an auditory prompt to bias listeners
- The auditory prompt was the carrier phrase "I really like" or "I weawwy yike" spoken by a single child.
  - The stimuli were CV sequences with initial /s/ or /θ/
    - In one experiment, listeners judged intermediate productions more leniently when they thought they were produced by younger speakers (Schellinger, 2008)
    - In another experiment, listeners judged intermediate productions more strictly when they thought they were produced by younger speakers (Munson et al., 2010)
    - The only difference between the two experiments was the instructions to the adult listeners.
  - A problem with the Schellinger et al. experiments was that the carrier phrase was produced by a single speaker, while the stimuli were produced by multiple speakers.

## PURPOSE OF THIS EXPERIMENT

- To determine whether an auditory prompt that was produced by the same child as the stimuli CV sequences biased listeners more consistently
- Prediction:
  - There would be a consistent effect of the auditory prompt
  - Based on previous research, we were not sure whether the auditory prompt would bias listeners to be more lenient or more strict in their accuracy judgments of the CV stimuli

## METHOD: EXPERIMENT 1

**Purpose:** To determine whether word lists composed of correctly or incorrectly produced words could affect listener judgments of speaker age.

**Participants:**

- 20 young adult native speakers of English

**Stimuli:**

- Word lists containing four words taken from recordings of 2- to 5-year-old typically developing English-speaking children
- A *correct* and an *incorrect* list was created with productions from each child speaker who contributed consonant vowel (CV) sequences to experiment 2.
  - *Correct* list: four words, each produced without errors
  - *Incorrect* list: four words, each produced with one or more errors

**Procedure:**

- Carrier word lists were presented in random order
- Participants clicked on a line to indicate their opinion on the age of the speaker
  - The left boundary indicated a very young child ("less than 2 years of age")
  - The right boundary indicated a somewhat older child ("more than 5 years of age")

**Analysis:**

- We analyzed the average age estimate for each child speaker based on: 1) actual age of child; and 2) whether the *correct* list or the *incorrect* list had been presented.

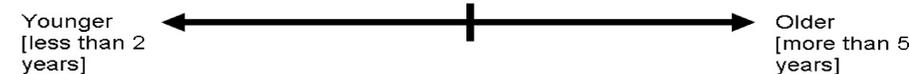


Figure 1. Rating display for experiment 1

## RESULTS: EXPERIMENT 1

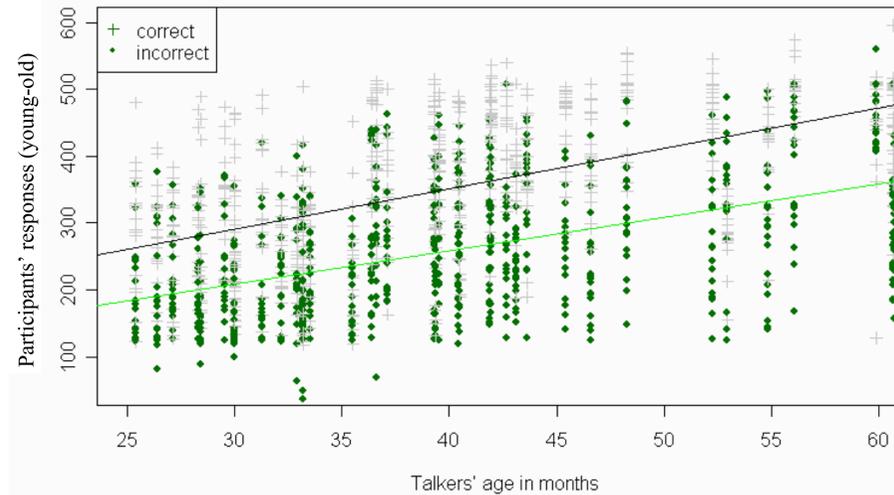


Figure 2. Scatter plot of estimated age as function of talkers' age in months

1. Listener judgments of a speaker's age increased as the talker's actual age increased.
2. The accuracy of the productions in the word lists affected the participants' judgments of age
  - *Correct* lists were judged as being produced by older children
  - *Incorrect* lists were judged as being produced by younger children.
3. This is illustrated in Figure 2: the regression line for the *correct* lists is above the regression line for the *incorrect* lists.

## METHOD: EXPERIMENT 2

**Purpose:** To determine if listeners' judgments of speech sound accuracy could be biased by their perception of speaker age.

- *Correct* list: Conveyed older child
- *Incorrect* list: Conveyed younger child

**Participants:**

- 20 young adult native speakers of English

**Stimuli:**

- Children's productions of consonant-vowel (CV) sequences containing word- initial /d/ or /g/ from the following transcription categories:
  - Correct productions of /d/ and /g/
  - Clear substitutions of [d] for /g/ and [g] for /d/
  - Tokens transcribed as intermediate between the target sounds /d/ and /g/
- *Correct* and *incorrect* lists from experiment 1.

**Procedure:**

- Participants heard each CV twice, once after the *correct* list and once after the *incorrect* list.
- Listeners participated in two blocks:
  - Block 1: "Was the "d" sound produced correctly?"
  - Block 2: "Was the "g" sound produced correctly?"

**Instructions to listener:**

- When children are learning to talk, they often make mistakes such as saying dood for good. In this experiment, you will be listening to many different children who are learning to produce "g" at the beginning of a word. Some of their productions will be correct and some will be mistakes such as dood for good. For each item, you will hear a list of four words produced by a child followed by a syllable that begins with /g/. Your job is to answer the question "Is it a correct /g/?" as quickly and accurately as possible.

**Analysis:**

- We compared the average accuracy score for each CV, based on: 1) transcription category; and 2) whether it was paired with the *correct* or *incorrect* list.

## RESULTS: EXPERIMENT 2

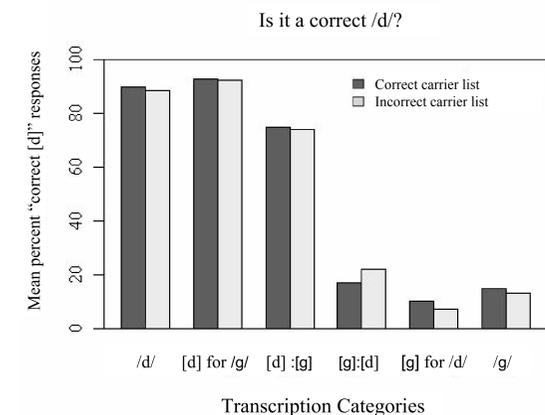
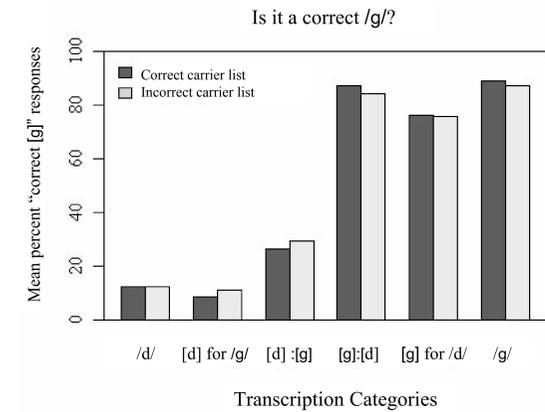


Figure 3. Responses to "Is it a correct /g/?" (top plot) and "Is it a correct /d/?" (bottom plot) as a function of transcription category and carrier list. Black bars show percent "yes" responses for stimuli preceded by "correct" carrier lists and gray bars for stimuli preceded by "incorrect" carrier lists.

- Results were similar for both the "Is it a correct /g/?" and the "Is it a correct /d/?" blocks
- For both blocks, there was a significant main effect of transcription category on percent "correct" responses.
- For both blocks, there was no significant main effect of carrier list and no significant transcription x list interaction.
- For both blocks, post-hoc comparisons showed that all transcriptions categories were significantly different from each other with the exception of:
  - No significant difference between correct /d/ and [d] for /g/ substitutions.
  - No significant difference between [g]:[d] and correct /g/

## CONCLUSION & DISCUSSION

- We were unable to bias listeners with an auditory prompt in this experiment.
- It is possible that we saw no effect of carrier list because listeners judge stop consonants more categorically than they judge fricatives, given the different results of Schellinger and colleagues.
- To test this hypothesis, we plan to try this same carrier list method with fricatives instead of stop consonants.
- More generally, these results, taken together with the results of Schellinger and colleagues, suggest that listener bias based on auditory prompts alone is relatively fragile and differs as a function of both instructions to listeners and perhaps, characteristics of the stimuli.

## ACKNOWLEDGMENTS

Supported by a Hilldale Undergraduate Research Fellowship to Amy Andrzejewski and NIDCD Grant R01DC02932 and NSF Grant BCS-0729140 to Jan Edwards.