



# Perception of Speech Categories in Musicians vs. Non-Musicians

Carly Faulhaber, Eun Jong Kong, Jan R. Edwards  
University of Wisconsin-Madison



## INTRODUCTION

- Speech and musical perception are influenced by musical training.
  - Musicians have significantly better musical pitch perception than non-musicians (Schon et al., 2004).
  - English speakers who are musicians can better discriminate lexical tone (a linguistic pitch pattern) in Mandarin Chinese than those who are non-musicians (Wong et al., 2006).
  - Musicians have enhanced periodicity encoding relative to non-musicians (Musacchia et al., 2007).
- In speech perception, some listeners focus primarily on phonetic categories, while others attend more to fine phonetic detail.
  - Kong & Edwards (2011) identified two kinds of listeners who accounted for about 50% of participants.
    - Categorical listeners: attended to primary acoustic cue only
    - Gradient listeners: attended to both primary and secondary cues.
- However, no studies have explored whether musicians are more likely to be gradient listeners than non-musicians.

## RESEARCH QUESTIONS

- To examine whether there is a difference between musicians' and non-musicians' perception of speech categories *not* related to pitch.
  - Are musicians more likely to be gradient listeners than non-musicians?
- To explore whether gradiency of perception is affected by consonant manner.
  - Predictions:
    - Musicians are more likely to be gradient listeners than non-musicians.
    - Fricatives are more likely to be perceived gradually than stops.

## METHOD

- Participants:**
  - 20 young adult musicians recruited from UW-Madison.
    - At least 8 years of continuous musical training
    - Practiced at least 10 hours per week
    - Currently involved in a performing ensemble
  - 20 young adult non-musicians recruited from UW-Madison
    - No more than 3 years of musical training
    - No musical training within the last 5 years
- Stimuli:** 400 children's production of consonant-vowel (CV) syllables paired with 5 vowels.
  - Use of child productions:
    - Natural stimuli preferred to synthetic stimuli.
    - Provide natural variation.
  - Set #1: 200 CV fricative syllables extracted from English words.
    - Half with /s/ and half with /θ/ (e.g., "sank" vs. "thank")
      - Correct productions of /s/ and /θ/ ("th")
      - Clear substitutions of [s] for /θ/ and [θ] for /s/
      - Tokens transcribed as intermediate between the target sounds /s/ and /θ/.
  - Set #2: 200 CV stop syllables extracted from English words.
    - Half with /t/ and half with /d/ (e.g., "tad" vs. "dad")
      - Various ranges of acoustic values between two sounds of /t/ and /d/
      - Selected based on acoustic properties (e.g., VOT).
- Procedure:** Participants used VAS (Visual Analog Scaling) on a laptop to rate each CV syllable. (See Figure 1)
  - Participants were asked to click a mouse on a line to indicate how close they perceived the sound to be to either endpoint.
  - Half of the participants in each group received the fricative stimuli first, and half received the stop stimuli first.
  - CV sequences were presented through headphones in a random order.

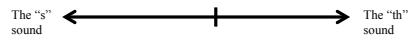


Figure 1: Visual Analog Scaling

## Analysis:

- We examined the location and the frequency of the clicks to differentiate gradient and categorical listeners.
- We compared the distributions of clicks between musicians and non-musicians and between the two sets of stimuli (fricatives vs. stops).
- See Figure 2 below for an example of a gradient listener vs. a categorical listener.

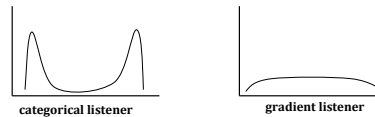


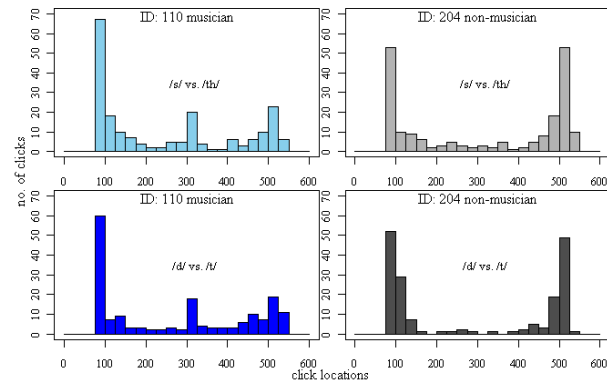
Figure 2: Example of Listeners

## RESULTS

### I. Individual characteristics

- Musicians vs. Non-musicians:** More musicians were gradient listeners while more non-musicians were categorical listeners.
  - /s/ vs. /θ/ :
    - 6 musicians and 3 non-musicians were judged as gradient listeners
    - 3 musicians and 7 non-musicians were judged as categorical listeners.
    - 11 musicians and 10 non-musicians did not fit clearly into either category.
  - /d/ vs. /t/ :
    - 4 musicians and 3 non-musicians were judged as gradient listeners
    - 8 musicians and 11 non-musicians were judged as categorical listeners.
    - 8 musicians and 6 non-musicians did not fit clearly into either category.
- Fricatives vs. Stops:** The fricatives (/s/ vs. /θ/) were perceived more gradually than the stops (/d/ vs. /t/). This was observed across both musicians and non-musicians.
  - /s/ vs. /θ/ :
    - 9 gradient listeners
    - 10 categorical listeners
    - 21 listener did not fit clearly into either category.
  - /d/ vs. /t/ :
    - 7 gradient listeners
    - 19 categorical listeners.
    - 14 listeners did not fit clearly into either category.

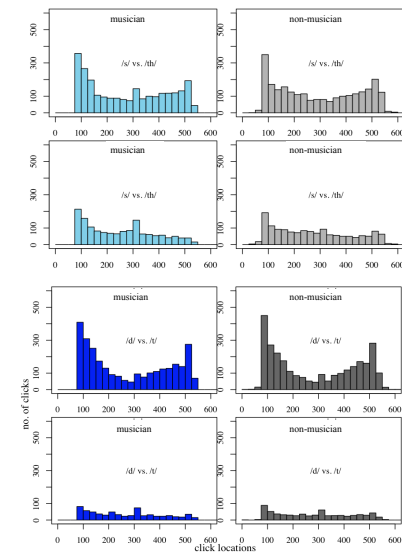
Figure 3 Histograms of click locations by a musician who perceived the sounds in a gradient manner and a non-musician who perceived the sounds in a more categorical manner. (x axis= location of clicks; y axis= frequency of clicks)



## II. Group characteristics

- The patterns observed in the individual listeners were also seen when the data were collapsed across listener group (see figure 4).
- The pattern was maintained even when the data was separated by transcription category ("correct" vs. "intermediate" productions).
  - For both productions that were coded as "correct" and those coded as "intermediate," non-musicians demonstrated a peak only at the endpoints of the VAS scale, while musicians displayed an additional peak in the middle.
  - This was observed for /s/ vs. /θ/ stimuli, but not for /d/ vs. /t/ stimuli.

Figure 4. A series of histograms of click locations for VAS for /s/ vs. /θ/ (top 2 plots) and /d/ vs. /t/ (bottom 2 plots) averaged across musicians (left plots, in blue) and non-musicians (right plots, in gray). For both sets of plots, responses to "correct" productions are shown on the top and responses to "intermediate" productions are shown on the bottom.



## DISCUSSION AND CONCLUSION

- We observed much individual variation in both groups of listeners.
- Nevertheless, more musicians were classified as gradient listeners than non-musicians.
- These results suggest that musical training may result in different speech perception abilities for musicians that are not simply related to lexical pitch perception.
- It was not surprising that fricatives were perceived more gradually than stops, given that the acoustic cues for fricatives are of less transient than those of stops.
- We have not yet been able to quantify our results.
  - A Chi-Square test was used but was not significant, given the small n (40) and the within-group variability.
  - We are currently exploring other statistical methods for comparing the histograms.
- We are also interested in exploring whether some of the variability within the musician group might be related to the instrument that the musician plays.
  - For example, are musicians who play a sting instrument more likely to be gradient listeners relative to musicians who play a wind instrument?

## ACKNOWLEDGMENTS

- Supported by a Hilldale Undergraduate Research Fellowship to Carly Faulhaber.
- Supported by NIH grant 02932 and NSF grant BCS 0729140 to Jan Edwards.