Expectation and Bias in the Assessment of Speech Naturalness

Megan L. Meyer, B.A.[1,2],
Benjamin Munson[1] [presenting author],
Anna Thurmes, M.A., CCC-SLP[3],
Kristy Benoit, M.A., CCC-SLP[4],
Kelly Nett Cordero, Ph.D., CCC-SLP[5],
& **Adriane L. Baylis**, Ph.D., CCC-SLP[6]

[1]Department of Speech-Language-Hearing Sciences, University of Minnesota, Minneapolis, MN,
[2]Department of Communicative Sciences and Disorders, New York University, New York, NY,
[3]School of Dentistry, University of Minnesota, Minneapolis, MN
[4]Lord Selkirk School Division, Selkirk, Manitoba, Canada,
[5]Gillette Children's Specialty Health Care, St. Paul, MN,
[6]Children's Hospital of Wisconsin, Milwaukee, WI

This presentation available at http://www.tc.umn.edu/~munso0005

De-Mystifying the Author List (De-List-ifying?)

- This presentation reports on the outcome of **Megan Meyer**'s undergraduate thesis, advised by **Benjamin Munson**. This was a follow up to a 2008 MA thesis by **Kristy Benoit** (also advised by Munson), which was an offshoot of a longitudinal project run by **Anna Thurmes**, **Kelly Nett Cordero**, and **Adriane Baylis**

Meyer, Munson, Thurmes, Benoit, Cordero, & Baylis, ASHA 2009

This presentation available at http://www.tc.umn.edu/~munso0005

The Gold Standard

- Regardless of the communication impairment we're working on, our goal should be for our clients' spoken messages to be maximally natural-sounding to the general public
- All of our objectives relating to specific behaviors (respiration, phonation, resonance, resonance, prosody, etc.) should be in the interest of facilitating natural-sounding speech
- "Natural" is a percept, a rating

What is Naturalness?

- It behooves us, then, to understand the nature of these ratings
  - Their psychometric properties (which tell us what kinds of scales should be used to elicit them)
  - Their relationship to specific speech behaviors (so that we know what to work on to achieve maximally natural speech)
  - Their susceptibility to bias

Today's Talk

Bias is Everywhere

- Speech perception is biased by linguistic factors (see Kent, 1996, for a review)
  - We build up detailed models of what to expect in an utterance, consequently,
  - We 'hear' things that aren't in the signal when we expect them to be there, and…
  - We interpret ambiguous things differently depending on how we think we should interpret them

Bias is Everywhere

- Perception is also biased by social factors
  - **Native Language** (Li, Munson, Beckman, Edwards, Yoneyama, & Hall, 2008)
  - **Regional dialect** (Hay, Warren, & Drager, 2006; Niedzielski, 1999)
  - **Age** (Drager, 2008)
  - **Gender** (Johnson, Strand & D'Imperio, 1999; Munson, 2009; Munson & Seppanen, 2009 [this conference])
  - **Race** (Staum Casasanto, 2008)
  - **Presence of disorder** (Munson, Edwards, Schellinger, Beckman, & Meyer, forthcoming; Schellinger, Edwards, Beckman, & Munson, 2008)

Today's Talk

Knowledge of Disorder: Children's Articulation

- Munson et al. (forthcoming) and Schellinger et al. (2008) found weak evidence that listeners change their criteria for /s/ accuracy when they are led to believe that the child they are listening to has a speech-sound disorder
- These effects can occur in both directions
- Schellinger (2008) found that these effects were equivalent for Communication Disorders undergraduate students, and graduate students in Speech-Language pathology

Knowledge of Disorder: Cleft Palate/Lip (CP±L)

- Podol and Salvia (1976) found ratings of nasality of a single passage were higher when this passage was presented concurrent with a picture of a child with repair-surgery scars than without
- Glass and Starr (1979) used a wider range of samples of the speech of people with CP±L and pictures. They failed to replicate Podol and Salvia, though they did find that nasality impacted ratings of speakers' attractiveness.

Knowledge of Disorder: Cleft Palate/Lip (CP±L)

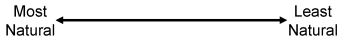
- Sinko and Hedrick (1982) found that ratings of speech acceptability of people with CP±L were not significantly different in audio-only and audio-visual conditions.
- Lallh and Putnam Rochet (2000) found that university undergraduates rated speakers with voice and resonance disorders more negatively than speakers without, and that this was not alleviated by giving them information about these disorders.

Benoit et al. (2008)

- Gathered measures of speech naturalness ("typical speech you would expect to hear in any given situation") from naïve listeners to use as outcome measures in a retrospective study of people with CP±L who had received services at the University of Minnesota Cleft Palate Clinic.
- Benoit's subjects were divided into two groups: those who were told that they speakers had a history CP±L given information about it, and those who were not.

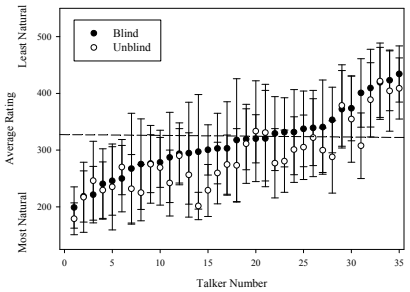
Benoit et al. (2008)

- Ratings were elicited using a visual-analog scale.



Benoit et al. (2008)

- There were subtle differences between the mean ratings for the group who knew the diagnosis ('unblind') and the group that did not ('blind')
- The 'unblind' group were more generous in their ratings



Benoit et al. (2008)

- Moreover, regression analysis showed that the ratings by the 'unblind' group were predicted more strongly by historic measures of the talkers' articulation ability than were the blind group's ratings.

The Current Study

- To examine whether ratings of naturalness differ as a function of whether the samples are paired with a picture of a person with or without obvious scarring from surgery to repair a cleft palate/lip
- If a difference is found, to examine whether the predictors of the ratings differ systematically (as did the blind and unblind groups in Benoit et al. 2008)

Speech Samples

- Thirty-five samples of a connected speech sample ('lazy Jack') with a history of cleft palate/lip, who had been seen at the University of Minnesota Cleft Palate/Craniofacial clinic
  - Identical to those from Benoit et al. (2008)

Faces

- High-quality digital pictures of individuals who had been seen at the University of Minnesota Cleft Palate/Craniofacial and Oral Surgery clinics.
- Included people with obvious scarring and malocclusions, and people whose scars were sufficiently posterior to not be visible externally

Voice-Face Pairing

- The voices were paired with one picture of a same-gender, (approximately) same-age picture. In one condition, the narrative was paired with a picture without obvious scarring; in the other condition, it was paired with a picture with obvious scarring.
- The speech samples were divided into two groups, such that each listener saw approximately equal numbers of scar and non-scar pictures. We call these groups *speech-sample groups*.
  - I.e., one group had narratives 1-17 paired with a scar picture, and narratives 18-35 paired with a non-scar picture. The other group had the opposite pairing
  - The two groups of talkers did not differ significantly in the ratings attributed to them by the listeners in Benoit et al. (2008).

Listeners

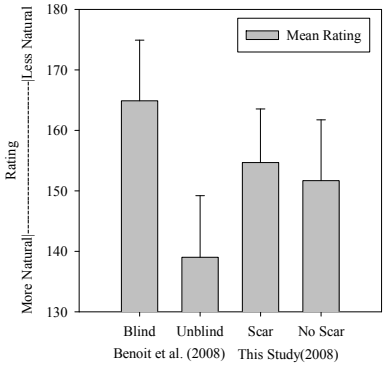
- Twenty members of the University of Minnesota community.
- The listeners were told that the pictures were the talkers who produced the narratives
  - This deception was debriefed following the experiment
- Ten listeners each participated in the two speech-sample groups.

Analysis

- Click location in pixels was used as the dependent measure
- A two-factor mixed-model ANOVA with face scarring as the within-subjects factor and speech-sample group as the between-subjects factor.
- Surprisingly, the only significant effect was one we planned not to get—of speech-sample group!
  - This appeared to be due to the increased power that the within-subjects comparison in this study offered.

Analysis

- Not surprisingly, regression models were similar for the two listener groups. They were similar to the regression models for Benoit et al.'s blind group
- Moreover, the ratings were intermediate between those for Benoit et al.'s



Why is Bias so Inconsistent?

- Conclusion: suggesting the talker's cleft status with pictures does not affect ratings of speech naturalness.
- But this contrasts with previous studies with the very same group, begging the question: *Why are bias effects so inconsistent?*
- The answer may be in how obviously it is suggested

Why is Bias so Inconsistent?

- The *overt mention* of disorder in Benoit et al. made listeners both more generous in their ratings, and made their ratings more closely related to characteristics of the speech samples
- The *covert suggestion* of disorder in this study had no effect.
- Naturalness may be a sufficiently abstract parameter that it is only susceptible to bias when that bias is introduced explicitly.
- This predicts that there would be a relationship between the automaticity with which a speech attribute is processed, and its susceptibility to bias from overt mention or tacit suggestion.

Acknowledgments

- Karlind Moller for input on earlier stages of this project, Stephanie Holzwart for digitizing the speech samples, and James Swift for providing the pictures, University of Minnesota College of Liberal Arts for funding this research, and Mary Kennedy for providing valuable feedback on the written version of Megan Meyer's thesis

References

Benoit, K., Munson, B., Thurmes, A., Cordero, K.N., Baylis, A., & Moller, K. (2008). *Factors Affecting Speech Naturalness in Young Adults with a History of Cleft Palate.* Poster presented at the 2008 ASHA Convention, Chicago, 20-22.

Drager, Katie (2006). From Bad to Bed: the relationship between perceived age and vowel perception in New Zealand English. *Te Reo, 48*, 55-68.

Glass, L., & Starr, C.D. (1979). A study of relationships between judgments of speech and appearance of patients with orofacial clefts. *Cleft Palate Journal, 16*, 436-440.

Hay, J., Warren, P., & Drager, K. (2006). Factors influencing speech perception in the context of a merger-in-progress. *Journal of Phonetics, 34*, 458-484

Johnson, K, Strand, E. & D'Imperio, M. (1999). Auditory-visual integration of talker gender in vowel perception. *Journal of Phonetics, 27*, 359-384.

Kent, R. (1996). Hearing and believing: some limits to the auditory-perceptual assessment of speech and voice disorders. *American Journal of Speech-Language Pathology, 5*, 7-23.

Lallh, A.K., & Putnam Rochet, A. (2000). The effect of information on listeners' attitudes toward speakers with voice or resonance disorders. *Journal of Speech, Language, and Hearing Research, 43*, 782-795.

Munson, B. (2009). *Gender biases in fricative identification revisited.* Oral presentation at the annual meeting of the Linguistic Society of America, San Francisco, CA.

Munson, B., Edwards, J., Schellinger, S.K., Beckman, M.E., & Meyer, M.K. (submitted). *Deconstructing Phonetic Transcription: Language-Specificity, Covert Contrast, Perceptual Bias, and an Extraterrestrial View of* Vox Humana. *Clinical Linguistics and Phonetics*

Munson, B., Li, F., Yoneyama, K., Hall, K., Beckman, M., Edwards, J., & Sunawatari, Y. (2008). *Sibilant Fricatives in English and Japanese: Different in Production or in Perception?* Oral presentation given at the annual meeting of the Linguistic Society of America, Chicago, IL.

Munson, B. & Seppanen, V. (2009). *Perceived gender affects ratings of the quality of children's spoken narratives.* Oral presentation given at the annual meeting of the American Speech-Language-Hearing Association, New Orleans, LA.

Niedzielski, N. (1999). The effect of social information on the perception of sociolinguistic variables. *Journal of Language and Social Psychology, 18*, 62-85.

Podol, J., & Salvia, J. (1976). Effects of visibility of a prepalatal cleft on the evaluation of speech. *Cleft Palate Journal, 13*, 361-366.

Schellinger, S., Edwards, J., Munson, B., & Beckman, M. E. (2008). *Assessment of children's speech production 1: Transcription categories and listener expectations.* Poster presented at the 2008 ASHA Convention, November 20-22.

Sinko, G.R., & Hedrick, D.L. (1982). The interrelationships between ratings of speech and facial acceptability in persons with cleft palate. *Journal of Speech and Hearing Research, 25*, 402-407.

Staum Casasanto, Laura. (2008). *Experimental Investigations of Sociolinguistic Knowledge.* Dissertation, Stanford University Department of Linguistics.

Meyer, Munson, Thurmes,
Benoit, Cordero, & Baylis,
ASHA 2009

This presentation available at
http://www.tc.umn.edu/~munso0005