

Implicit Processing, Social Stereotypes, and the 'Gay Lisp'

Sara Mack¹ and Benjamin Munson²

University of Minnesota

¹Department of Spanish and Portuguese Studies

²Department of Speech-Language-Hearing Sciences

Linguistic Society of America Annual Meeting

January 6, 2008

Background

- Socio-indexicality: Phonetic variation indexes different social categories
- Speech and sexual orientation
 - Empirical evidence
 - Stereotypes: the 'gay lisp'

Background

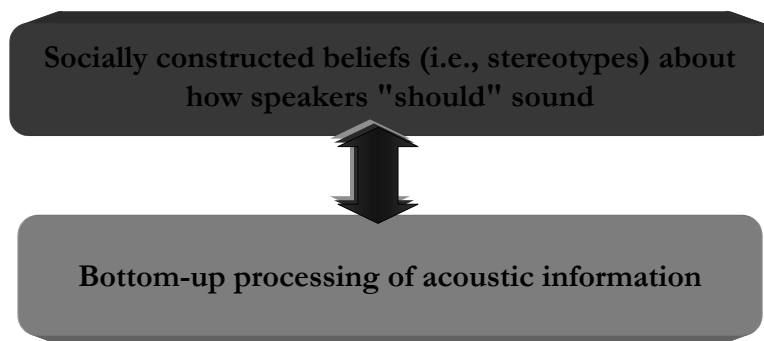
- Previous studies: explicit measures
 - Listener evaluations gathered via binary choice tasks, ratings scales, etc.
 - Assumes sequential processing of acoustic information:



- Potential drawbacks:
 - Evidence from social psychology of impact of “social desirability concerns” (Fazio & Olson, 2003) on responses
 - Impact of stereotypes on speech processing

Background

- Stereotypes and speech processing: the case for implicit measures



(Diagram based on Strand, 1999, p. 87.)

Goals of the Study

- *General goal:* To develop and test an implicit processing paradigm to explore listeners' sensitivity to associations between social categories and pronunciation variation
- *Specific focus:* to examine the interaction between perceptual characteristics of word-initial /s/ and perceived speaker sexual orientation on recognition accuracy for voices

Research questions

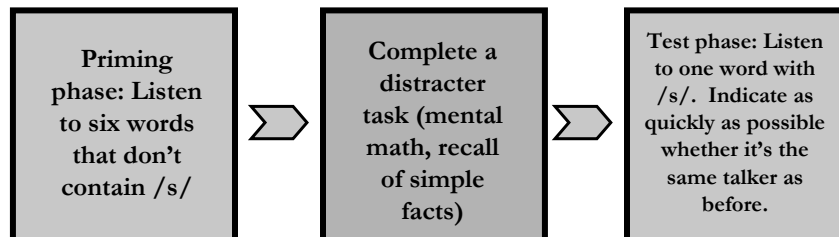
- *Question 1.* Do social stereotypes about fricative production affect voice recognition?
 - 1.A. Are gay-sounding talkers' voices recognized more quickly when their productions include /s/ variants that are stereotypically associated with gay-sounding speech?
 - 1.B. Are straight-sounding talkers' voices recognized more quickly when their productions include /s/ variants associated with heterosexual-sounding speech?
- *Question 2.* Is listeners' memory for voices in an implicit recognition task related to independently made measures of their perceived sexual orientation?

Method

- Immediately Primed Voice-Recognition Task
 - Measures the effect of exposure to a voice in a ‘priming’ phase to recognition accuracy and speed in a ‘test’ phase
 - Allows for a measure of relationships between variables without the overt mention of a social category
- Listener participants
 - 20 total (17 female, 3 male)
 - Native speakers of North American English
 - Mean age 29.1 years

Method

- Each experimental trial consisted of three phases:



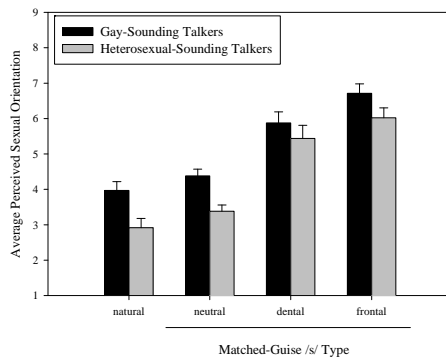
Stimuli

- Six priming phase words with no /s/ (*bell, fade, path, hoop, note, tooth*).
- Two test phase words (*sad, soon*), each with naturally produced and digitally manipulated tokens, resulting in four variants:
 - One unedited token with the naturally produced /s/
 - Three tokens with /s/ produced by a trained phonetician, appended to naturally produced VC bases:
 - Neutral /s/
 - Mimicked the characteristics of the average /s/ in this group
 - Dental /s/
 - Frontal /s/

Stimuli

- Produced by 16 men whose perceived sexual orientation (PSO) had been reported in a previous study (Munson et al., 2006)
 - Included the eight men with the most gay-sounding voices and the eight men with the most heterosexual-sounding voices from a group of 22 men, based on the ratings of the six prime words
 - PSO had been measured previously in *numerous* experiments, as had other characteristics (perceived height, perceived speech clarity, perceived age)
- Dialectally uniform, late teens through early thirties

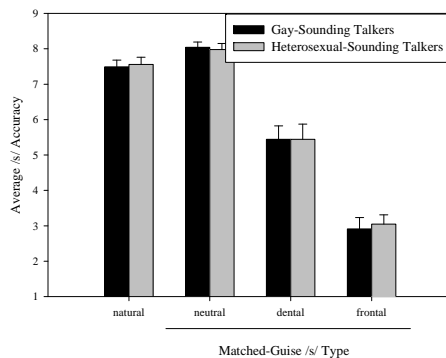
Stimuli



Graph 1. Average talker perceived sexual orientation ratings for the three matched-guise fricatives and the natural unedited stimuli, separated by talker perceived sexual orientation, from Munson and Zimmerman (2006).

- Interactions between perceived sexual orientation (PSO) ratings and /s/ variants for the target stimuli were analyzed by Munson and Zimmerman (2006)
- Dental and frontal /s/ made all speakers sound gayer, and reduced the PSO differences between the groups

Stimuli



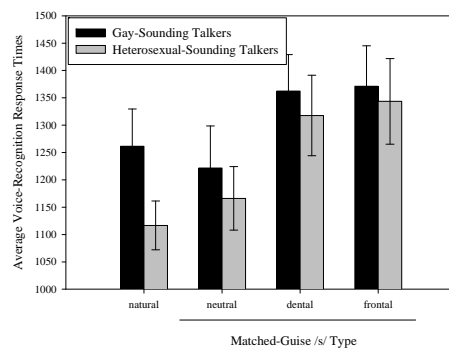
Graph 2. Average ratings of /s/ accuracy for the three matched-guise fricatives and the natural unedited stimuli, separated by talker perceived sexual orientation, from Munson and Zimmerman (2006).

- The dental and frontal /s/ were rated to sound less accurate than the natural and neutral /s/
- This *did not* interact with the PSO of the talker to which the /s/ was appended

Results

- Analysis based on response times (RT's) of “hits”; high rate of overall response accuracy (> 85%)
- Analysis included responses that were:
 - 4000 ms or less
 - More than three standard deviations from the mean

Results



Graph 3. Average listener response times for the three matched-guise fricatives and the natural unedited stimuli, separated by talker perceived sexual orientation.

- Significant main effect of /s/ type
 - Longer response times for the frontal and dental /s/ variants
- Significant main effect of perceived sexual orientation
 - Longer response times for the talkers who had been rated as more gay-sounding
- No significant interaction

Regression

- Regression analyses with standardized response times were used to examine:
 - 1. The possible influence of stimulus duration on RTs
 - Stimulus duration was not controlled
 - 2. The acoustic predictors of response times
 - 3. Whether response times were predicted by measures of PSO and production accuracy made by independent groups of listeners

Regression 1: Stimulus Duration

- Hierarchical Regression with Dummy-Coding
 - Dependent measure: average standardized response time
 - Step 1: stimulus duration
 - Step 2: dummy-code for /s/ type
 - Step 3: dummy-code for talker perceived sexual orientation
- Even when stimulus duration was accounted for, the following contrasts were significant in the full regression model:
 - {natural /s/, neutral /s/} vs. {frontal /s/, dental /s/}
 - Gay-sounding vs. Straight-Sounding
 - Stimulus duration was not significant in the full model

Regression 2: Acoustic Predictors

- Stepwise Multiple Regression
 - Dependent measure: average standardized RT
 - Independent measures: Fricatives' first and third spectral moments, vowels' f0, f1, and f2, stimulus duration
 - Each stimulus type examined separately
- Results
 - For the natural stimuli, f0 and f2 predicted variance in the RTs. Fricative acoustics did not
 - For the neutral /s/ stimuli, f0 predicted RTs
 - No predictors for the frontal and dental /s/
- Interpretation: /s/ characteristics do not affect RTs unless the /s/ is atypical.

Regression 3: Other Perceptual Measures

- Hierarchical Multiple Regression
 - Dependent measure: average standardized RT
 - Independent measures: stimulus duration, previously measured perceived sexual orientation (from Munson & Zimmerman, 2006), perceived accuracy
- Perceived sexual orientation predicted response times in a variety of regression models beyond what was predicted by measures of stimulus accuracy

Summary: Regression

- 1. The RT differences in the ANOVA are not an artifact of stimulus duration
- 2. For natural stimuli, vocalic characteristics affect RTs more strongly than /s/ characteristics
- 3. For all stimuli, perceived sexual orientation measures predict RTs better than measures of /s/ accuracy

Discussion and conclusion

- Social categories and pronunciation variation: speech processing is sensitive to associations between social categories and pronunciation variation
 - Previously-established perceived sexual orientation predicted response times
 - Provides further evidence of the impact of unconscious processes in linguistic processing of social information; supports Strand's (1999) claim that social information is stored in memory and is used in low-level, automatic speech processing
- A potential limitation: we were pairing an authentic cue (those in the VC) with a stereotypic cue (the lisped /s/). Maybe the priming only works when both cues are stereotypic.

Discussion and conclusion

- Perceived sexual orientation and /s/ type
 - Regression analysis suggest that processing speed in the voice recognition task was most closely related to the perceived sexual orientation of the talkers' voices, and not to perception of production accuracy
- Social information and the lexicon
 - Detail
 - Storage
 - Access

Acknowledgments

- Funded by...
 - The McKnight Foundation, through a University of Minnesota Presidential Fellowship to the second author
 - The Graduate School, University of Minnesota, through a grant-in-aid of research, scholarship, and creativity to the second author