

On Voiceless Fricative Perception: Vocal-Tract Normalization and Socioindexicality

Benjamin Munson
Department of Speech-Language-Hearing Sciences
University of Minnesota, Minneapolis, USA

Phonetics and Phonology Forum, Kobe, Japan, August 26, 2009

Benjamin Munson 1

Organization of talk

1. A model of phonological knowledge, and why sibilant fricatives are important
2. Vocal-tract normalization during voiceless sibilant fricative perception (Munson & Coyne)
3. Perceived gender and sibilant fricative perception (Munson & Kemper)

This talk is just one part of a larger research program that looks at production and perception across the lifespan, some of which I considered including in this talk. Please visit my website if you're interested. It's just not worth including it and then forcing myself to rush through this presentation.

Phonetics and Phonology Forum, Kobe, Japan, August 26, 2009

Benjamin Munson 2

A Model of Phonological Knowledge

- Laboratory phonology uses experimental and corpus-based methods to examine the cognitive representation of sound structure
- Basic observation 1: phonological categories do not have invariant articulatory or auditory correlates.

Phonetics and Phonology Forum, Kobe, Japan, August 26, 2009

Benjamin Munson 3

A Model of Phonological Knowledge

- Basic observation 2: Categories exist. People treat phonetically distinct instances of a sound (say, /s/) the same when asked to categorize them as either /s/ or some other phoneme.
- In other tasks, they retain the phonetic information
 - People can infer vowel contexts from sounds that have been excised from different vowel contexts (Soli, 1981)
 - People can identify attributes of speakers from phonetic detail (e.g., Munson, McDonald, DeBoe, & White, 2006, *inter alia*)
- Hence, knowledge of the sound structure of language involves knowledge of the parametric phonetic characteristics of sounds, and knowledge at multiple levels of abstraction away from these raw sensory experiences

Phonetics and Phonology Forum, Kobe, Japan, August 26, 2009

Benjamin Munson 4

A Model of Phonological Knowledge

- Understanding categories, then, involves both understanding the information that people encode, and the process that allows them to impute the existence of categories.
- This talk is about people's knowledge of voiceless fricatives.
- Why fricatives?
 - They're not vowels (which are pretty well studied and well understood)
 - (My assessment): they're the obstruent consonants whose articulation, acoustics, and perception are most widely studied
 - You could argue with me on this one
 - They are the locus of lots of socioindexical variation

Phonetics and Phonology Forum, Kobe, Japan, August 26, 2009

Benjamin Munson 5

Study 1: Vocal-Tract Normalization during Fricative Perception

Collaborator: Alexander Coyne

Phonetics and Phonology Forum, Kobe, Japan, August 26, 2009

Benjamin Munson 6

Vocal-Tract Normalization in Fricative Perception

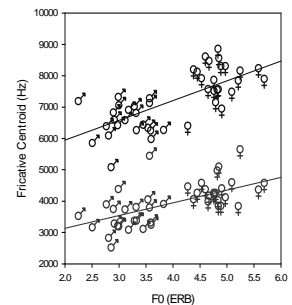
- During perception listeners must make inferences about sources of variation in speech so that they can associate variable acoustic signals with talkers' intended messages (see review in Johnson, 1990)
 - The phonetics 101 term is 'normalization'
- One well-studied source of variation: individual differences in vocal-tract morphology (see review in Johnson, 2006)

Phonetics and Phonology Forum, Kobe, Japan, August 26, 2009

Benjamin Munson 7

Vocal-Tract Normalization in Fricative Perception

- Fricative acoustics correlate with variables known to correlate with vocal-tract length
- Munson, McDonald, DeBoe, and White (2006): the centroid frequencies of /s/ (in blue) and /ʃ/ (in red) correlate with f_0 , which is correlated with vocal-tract length
- These are 44 talkers, 22 men, 22 women, intentionally over-sampling voices that are not prototypically male or female

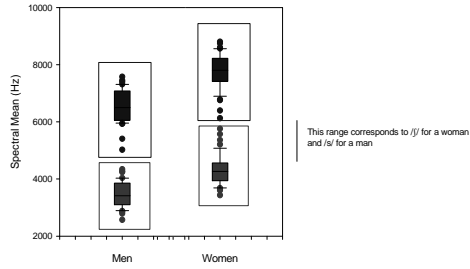


Phonetics and Phonology Forum, Kobe, Japan, August 26, 2009

Benjamin Munson 8

Vocal-Tract Normalization in Fricative Perception

- The net result: men and women produce sibilant fricatives with different spectral ranges. (/s/ is blue, /ʃ/ is red.) These data are from Munson et al. (2006).

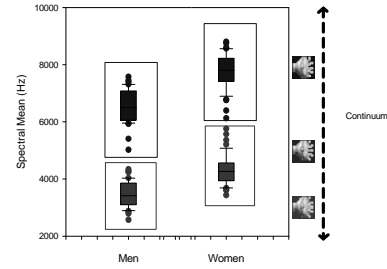


Phonetics and Phonology Forum, Kobe, Japan, August 26, 2009

Benjamin Munson 9

Vocal-Tract Normalization in Fricative Perception

- If you make a fricative continuum between the highest and the lowest frequencies and combine it with vowels produced by men and women, there should be some fricatives that are identified differently depending on the sex of the voice with which it is combined



Phonetics and Phonology Forum, Kobe, Japan, August 26, 2009

Benjamin Munson 10

Vocal-Tract Normalization in Fricative Perception

- ...and indeed there are.
- ...and the effect is gradient.
- Munson, Jefferson, and McDonald (2006, JASA): *sack-shack*

- 12% sack
- 25% sack
- 53% sack
- 80% sack

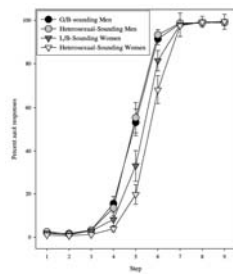


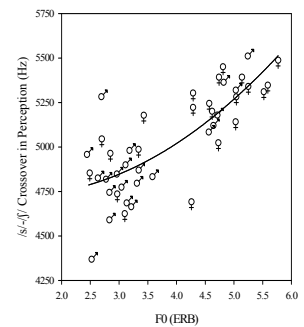
FIG. 1. Percent sack responses for GB and International Standard men and women on the nine-step sack-shack continuum. Error bars represent one standard error of measurement.

Phonetics and Phonology Forum, Kobe, Japan, August 26, 2009

Benjamin Munson 11

Vocal-Tract Normalization in Fricative Perception

- Might this just represent vocal-tract normalization?
- After all, crossover points for the 44 talkers in Munson, Jefferson, and McDonald (2006) are correlated with f_0 .
- ...but the correlation is only significant for the 22 women talkers, not for the male talkers.
- ...and the crossover points were most strongly correlated with perceptual measures of the sex typicality of talkers' voices rather than with perceptual or acoustic measures of talker size.



Phonetics and Phonology Forum, Kobe, Japan, August 26, 2009

Benjamin Munson 12

Vocal-Tract Normalization in Fricative Perception

- Alexander Coyne's 2007 *Summa cum Laude* BA thesis from the University of Minnesota asked...
- Do listeners normalize for differences in apparent vocal-tract size similarly when they're listening to men's and women's voices?
 - i.e., can we replicate Munson, Jefferson, and McDonald's finding manipulating only those parameters related most closely to vocal-tract size?
 - If listeners respond differently to acoustic manipulations in apparent vocal-tract size for men's and women's voices, then we have further evidence that normalization involves more than just making inferences about talkers' size.

Phonetics and Phonology Forum, Kobe, Japan, August 26, 2009

Benjamin Munson 13

Vocal-Tract Normalization in Fricative Perception

- Do listeners respond differently depending on what they are led to believe the source of the variation is?
 - If listeners respond differently depending on why they think talkers' voices differ, then again, we have evidence that normalization involves more than just making inferences about talkers' size.

Phonetics and Phonology Forum, Kobe, Japan, August 26, 2009

Benjamin Munson 14

Vocal-Tract Normalization in Fricative Perception

- Thirty-nine listeners identified 20 continua, 10 /s/-/ʃ/ and 10 /s/-/θ/ continua.
- /s/-/θ/ continua were included because /θ/ acoustics do not differ as a function of vocal-tract length.
- 10 were created by combining fricative continua with a man's production of /aɪ/ (from a natural production of the word *sigh*), 10 with a woman's production.
- Each man and each woman's production had their formant frequencies and f0 scaled to give five different apparent vocal-tract lengths.



Phonetics and Phonology Forum, Kobe, Japan, August 26, 2009

Benjamin Munson 15

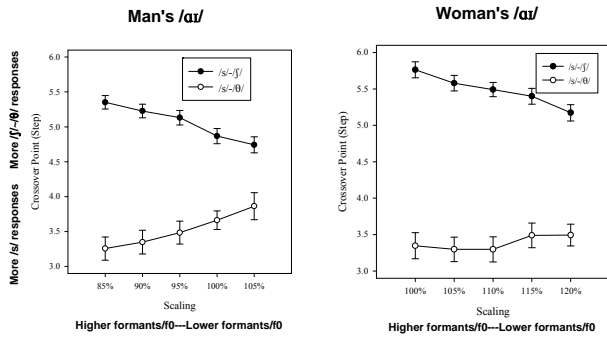
Vocal-Tract Normalization in Fricative Perception

- Thirteen listeners were not told anything about the stimuli; 13 were told that the stimuli were modeled on productions of people who differed in age; 13 were told that the stimuli were modeled on the productions of people who differed in height
 - The *height* and *age* groups did a second task in which they rated the age and height of the talkers for each stimulus.
- Probit slope and crossover points for each continuum were calculated.

Phonetics and Phonology Forum, Kobe, Japan, August 26, 2009

Benjamin Munson 16

Vocal-Tract Normalization in Fricative Perception



Phonetics and Phonology Forum, Kobe, Japan, August 26, 2009

Benjamin Munson 17

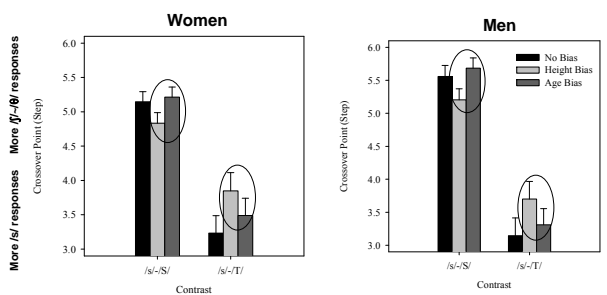
Vocal-Tract Normalization in Fricative Perception

- Scaling affected /s/-/ʃ/ perception equally for the man and the woman's /aɪ/.
 - Consistent with a vocal-tract normalization explanation
- Scaling affected perception of /s/-/θ/ only for the man's /aɪ/.
 - Not consistent with a vocal-tract normalization explanation
- There were differences in the three groups' (no bias, age bias, height bias) crossover points, when pooled across the five scaling conditions.

Phonetics and Phonology Forum, Kobe, Japan, August 26, 2009

Benjamin Munson 18

Vocal-Tract Normalization in Fricative Perception



Phonetics and Phonology Forum, Kobe, Japan, August 26, 2009

Benjamin Munson 19

Vocal-Tract Normalization in Fricative Perception

- Averaged across scaling conditions, listeners in the age and height bias groups did behave differently.
- People who were in the age-bias group were more likely to respond to /s/ in the /s/-/θ/ continua and /ʃ/ in the /s/-/ʃ/ continua.
 - They responded as if the intermediate tokens were English-appropriate developmental errors.
 - If they thought they were listening to people who differed in age, they accepted something intermediate between /s/ and /θ/ as /s/, and something intermediate between /s/ and /ʃ/ as /ʃ/.

Phonetics and Phonology Forum, Kobe, Japan, August 26, 2009

Benjamin Munson 20

Vocal-Tract Normalization in Fricative Perception

- The no bias group performed more similarly to the age-bias group than to the height-bias group
 - The most neutral interpretation of the tokens is that the people differ in age rather than in height.

Phonetics and Phonology Forum, Kobe, Japan, August 26, 2009

Benjamin Munson 21

Vocal-Tract Normalization in Fricative Perception

- Conclusion from this section: People recruit expectations about sources of variation when they are perceiving fricatives. Perception cannot be reduced entirely to peripheral 'vocal-tract normalization' processes.

Phonetics and Phonology Forum, Kobe, Japan, August 26, 2009

Benjamin Munson 22

Study 2: Perceived Gender and Fricative Perception

Collaborator: Sara Kemper

Phonetics and Phonology Forum, Kobe, Japan, August 26, 2009

Benjamin Munson 23

Perceived Gender and Fricative Perception

- Men and women produce fricatives differently. Fricative production is the locus of numerous popular-culture stereotypes about sex typicality and speech
- Strand and Johnson (1996) found different crossover points for synthetic fricatives combined with non-prototypical voices paired with male and female faces.

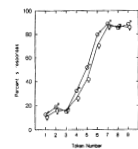


Figure 2.2. Results of Experiment 2. The effect of the gender of a visually-presented face on the perception of synthetic fricatives for the words produced by the non-prototypical male speaker.

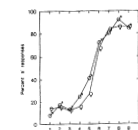


Figure 2.3. Results of Experiment 3. The effect of the gender of a visually-presented face on the perception of synthetic fricatives for the words produced by the non-prototypical female speaker.

Phonetics and Phonology Forum, Kobe, Japan, August 26, 2009

Benjamin Munson 24

Perceived Gender and Fricative Perception

- The audiovisual results suggest that perception is not just the compensation for differences in perceived vocal-tract size.
 - Later replicated (in spirit) with perception of an [ʌ]-[ʊ] continuum
- One interpretation: visual images activate a set of stored exemplars of instances of women producing fricatives. These bias responses to the ambiguous fricatives.
 - Consistent with explanations for the influence of other social variables on phoneme labeling (Drager, 2006; Hay, Warren, & Drager, 2006, Niedzielski, 1999)

Phonetics and Phonology Forum, Kobe, Japan, August 26, 2009

Benjamin Munson 25

Perceived Gender and Fricative Perception

- The goals of this experiment were to...
- **1. Replicate the /s/-/ʃ/ audio-visual finding**
- Replication is, quite simply, the foundation of good theory-building, period.
- **2. Examine the influence on gender with stimuli whose acoustic characteristics are consistent with similar-sized vocal tracts.**
 - The stimuli that Strand and Johnson and Munson, Jefferson, and McDonald used varied in the acoustic parameters that are known to be correlated with vocal-tract length, F3 and f0
 - This study examined the perception of men's and women's tokens that had been acoustically modified so that these parameters are equivalent.

Phonetics and Phonology Forum, Kobe, Japan, August 26, 2009

Benjamin Munson 26



Perceived Gender and Fricative Perception

- **3. Examine the influence of gender on perception of an /s/-/θ/ continuum**
- There isn't clear evidence that /θ/ acoustics are related to either sex or to vocal-tract length
- There are, however, ample social stereotypes about /s/ variation and gender typicality (i.e., Mack & Munson, 2008; Munson & Zimmerman, 2006)
- A finding that talker gender affects /s/-/θ/ perception would buttress the argument that these effects are not due to vocal-tract normalization.

Phonetics and Phonology Forum, Kobe, Japan, August 26, 2009

Benjamin Munson 27

Perceived Gender and Fricative Perception





- Eight continua: four *sigh-shy* continua and four *sigh-thigh* continua
- Created by appending fricative continua to two adults' natural productions of [aɪ] from the word *sigh* (originally used in Coyne, 2007)
 - One man: 
 - One woman: 

Phonetics and Phonology Forum, Kobe, Japan, August 26, 2009

Benjamin Munson 28

Perceived Gender and Fricative Perception

- Each base had its third-formant frequency (taken from the nucleus of the /a/ portion of the /aɪ/ diphthong) scaled to be consistent with either a 14 cm (F3 ≈ 3050 Hz) or a 17 cm (F3 ≈ 2525 Hz) vocal tract
- F0 was always ≈ 130 Hz
- We call this manipulation *Apparent Vocal-Tract Length* (aVTL)

	Man's [aɪ]	Woman's [aɪ]
14 cm		
17 cm		

Phonetics and Phonology Forum, Kobe, Japan, August 26, 2009

Benjamin Munson 29

Perceived Gender and Fricative Perception

- Bases were appended to two continua
- /s/-/ʃ/: The same continuum as Strand and Johnson, with the most /s/-like and the most /ʃ/-like stimuli removed
- /s/-/θ/: A seven-step continuum created by mixing the /s/ endpoint from the continuum above with a naturally produced /θ/ (following McGuire, 2007)
 - Relative intensity of the fricative and the vowel was varied across the continuum



Phonetics and Phonology Forum, Kobe, Japan, August 26, 2009

Benjamin Munson 30

Perceived Gender and Fricative Perception

- Each stimulus was presented four times with the same man's picture, four times with the same woman's picture, and four times with no picture
- We call this manipulation *Face-Bias*
- (Pictures taken from the Caltech Frontal Facial Database, http://www.vision.caltech.edu/Image_Datasets/faces/README)



Phonetics and Phonology Forum, Kobe, Japan, August 26, 2009

Benjamin Munson 31

Perceived Gender and Fricative Perception

- 19 Adults from the North-Central dialect region
 - Native speakers of English
 - No past or current speech, language, or hearing impairment

Phonetics and Phonology Forum, Kobe, Japan, August 26, 2009

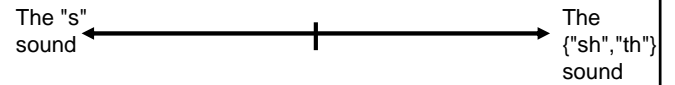
Benjamin Munson 32

Perceived Gender and Fricative Perception

- Stimuli were presented in a fully randomized design (*sigh* and *thigh* trials were intermingled, as were face-bias conditions and aVTL conditions).
- Listeners responded by clicking on a visual analog scale (VAS), following Urberg-Carlson, Kaiser, and Munson (2008)
- Click location (in pixels) was logged automatically

Phonetics and Phonology Forum, Kobe, Japan, August 26, 2009

Benjamin Munson 33



Phonetics and Phonology Forum, Kobe, Japan, August 26, 2009

Benjamin Munson 34

Perceived Gender and Fricative Perception

- Listeners were instructed to choose a click location that corresponded to their perception of the 'proximity' of the stimulus to a good "s," "sh," or "th" sound.

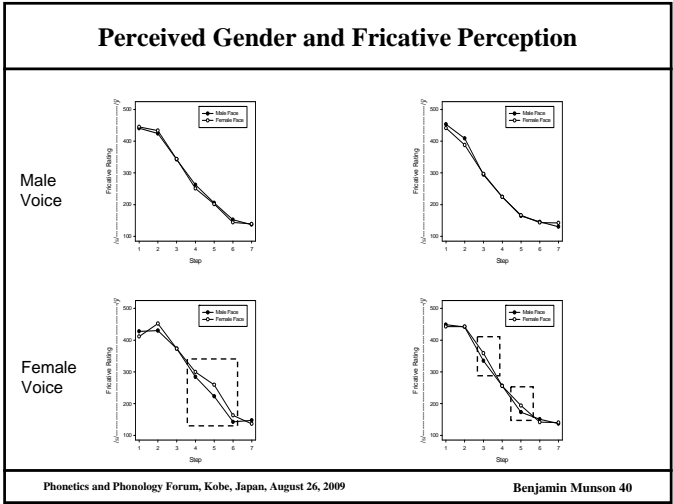
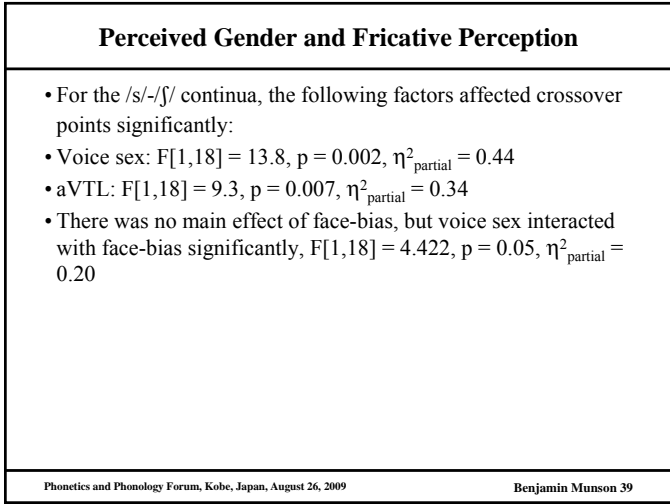
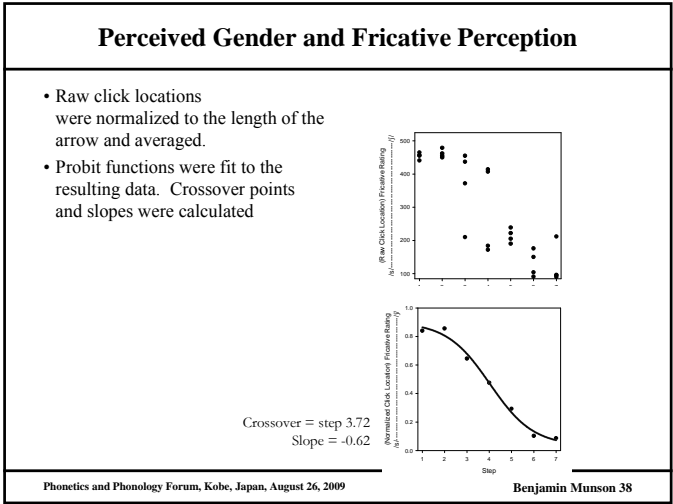
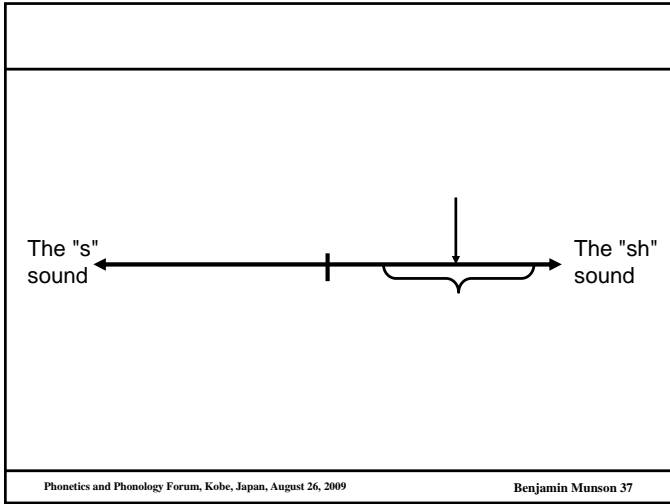
Phonetics and Phonology Forum, Kobe, Japan, August 26, 2009

Benjamin Munson 35



Phonetics and Phonology Forum, Kobe, Japan, August 26, 2009

Benjamin Munson 36



Perceived Gender and Fricative Perception

- The two stimuli that were most susceptible to face-biasing were...



Phonetics and Phonology Forum, Kobe, Japan, August 26, 2009

Benjamin Munson 41

Perceived Gender and Fricative Perception

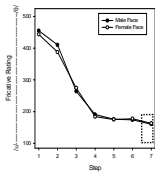
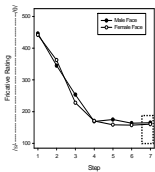
- For the /s/-/θ/ continua, the following factors affected crossover points significantly:
 - Voice sex: $F[1,17] = 7.2, p = 0.02, \eta^2_{\text{partial}} = 0.30$
 - aVTL: $F[1,17] = 29.3, p < 0.001, \eta^2_{\text{partial}} = 0.63$
 - Face-bias: $F[1,17] = 6.7, p = 0.02, \eta^2_{\text{partial}} = 0.28$
- None of the factors interacted

Phonetics and Phonology Forum, Kobe, Japan, August 26, 2009

Benjamin Munson 42

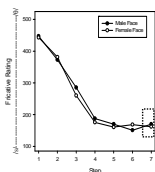
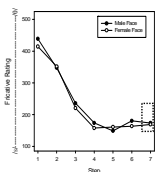
Perceived Gender and Fricative Perception

Male Voice



The effect of face-bias is small, but is very consistent across listeners

Female Voice

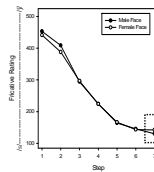
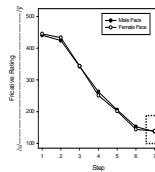


Phonetics and Phonology Forum, Kobe, Japan, August 26, 2009

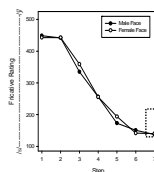
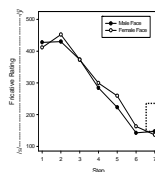
Benjamin Munson 43

Perceived Gender and Fricative Perception

Male Voice



Female Voice



Phonetics and Phonology Forum, Kobe, Japan, August 26, 2009

Benjamin Munson 44

Perceived Gender and Fricative Perception

- We found that the influence of voice gender on labeling is (statistically) independent from the influence of apparent vocal-tract size. It affects perception of at least two fricative continua
- More evidence that Strand and Johnson's findings indeed reflect listeners' perception of talker gender and not talker vocal-tract size.

Phonetics and Phonology Forum, Kobe, Japan, August 26, 2009

Benjamin Munson 45

Perceived Gender and Fricative Perception

- Why does voice gender have a *stronger* influence on labeling than does apparent vocal-tract length for an /s/-/ʃ/ continuum, but a *weaker* influence than apparent vocal-tract length for perception of an /s/-/θ/ continuum.
- How do listeners interpret the acoustic manipulations of apparent vocal-tract length?
 - /s/-/ʃ/: vocal-tract length
 - /s/-/θ/: age

Phonetics and Phonology Forum, Kobe, Japan, August 26, 2009

Benjamin Munson 46

Perceived Gender and Fricative Perception

- The problem with Munson (2009) (and Strand and Johnson, 1996, and many others): mentioning gender explicitly runs the risk of invoking stereotypes about men's and women's fricatives, rather than actual experiences.
- An alternative: *imply* the gender, and see if this affects fricative perception
- In her 2009 BA thesis from Vassar College, Sara Kemper examined this.

Phonetics and Phonology Forum, Kobe, Japan, August 26, 2009

Benjamin Munson 47

Perceived Gender and Fricative Perception

- The method was borrowed (broadly) from studies by Eisner, McQueen, and Cutler.
 - Phase 1: Identify an /s/-/ʃ/ continuum with a gender-neutral voice
 - Phase 2: Do a task that suggests the gender of the talker
 - One group is led to believe that the talker is a man, the other to believe that it's a woman
 - Phase 3: Identify the /s/-/ʃ/ continuum again
- Gender is never mentioned during the experiment.

Phonetics and Phonology Forum, Kobe, Japan, August 26, 2009

Benjamin Munson 48

Perceived Gender and Fricative Perception

- In Kemper's experiment, gender was suggested by presenting listeners with sentences that were judged by an independent group of listeners to be either very likely to have been produced by a man or by a woman. 🗣️ 🗣️
- The written sentences were judged to be either very likely to have been produced by a man or by a women
- A separate set of content-neutral sentences was used to determine the formant- and f0-scaling needed for the sentences to be perceived as neutral.

Phonetics and Phonology Forum, Kobe, Japan, August 26, 2009

Benjamin Munson 49

Perceived Gender and Fricative Perception

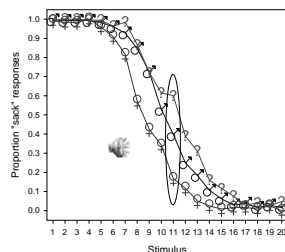
- Equal numbers of men and women participated, and..
- ...Simply put, the experiment didn't work!
- Identification patterns didn't differ for the group that was biased to believe that the talker was male and those biased to believe it was female.
- However, after the experiment, the 32 listeners were asked what they thought the sex of the talker was.
- 6 said "woman," 17 said "man," and 9 were uncertain, or said that their percept changed over the course of the experiment.

Phonetics and Phonology Forum, Kobe, Japan, August 26, 2009

Benjamin Munson 50

Perceived Gender and Fricative Perception

- Differences in fricative perception occur for gender-neutral stimuli as a function of the gender that the participants thought the talkers were.
- In the expected direction: those who thought they were listening to a woman heard more /ʃ/ tokens.



Phonetics and Phonology Forum, Kobe, Japan, August 26, 2009

Benjamin Munson 51

Perceived Gender and Fricative Perception

- There is evidence, that speakers recruit stereotypes about gender when perceiving fricatives, that these occur even when gender isn't mentioned explicitly, and that these have effects that are independent from effects of vocal-tract size.

Phonetics and Phonology Forum, Kobe, Japan, August 26, 2009

Benjamin Munson 52

The Broader Endeavor

- This finding and others cited herein suggest that speech perception is not merely perception of simple chains of deterministic events that are equivalent across talkers, contexts, and languages
 - The objects of speech perception are *not* IPA-style phonetic symbols!
- Rather, the process of speech perception references elements in a complex topography that involves representations and structures in at least three sensory domains (visual, auditory, tactile), and which evolve in individual speakers and hearers over a lifetime of perceiving different words in different contexts spoken by different talkers

Acknowledgments

- **Collaborators on research projects in this talk:** Mary Beckman, Alexander Coyne, Jan Edwards, Eden Kaiser, Fangfang Li, Sarah Kemper, Kari Urberg-Carlson
- **Funding:** McKnight Foundation, University of Minnesota College of Liberal Arts, and NSF grant BCS-0729277 to Benjamin Munson