Gender Differences in Speech: Within-session stability of acoustic features of conversational and clear speech for male and female talkers

Sarah Hargus Ferguson1, Shae D. Morgan1, Lydia R. Rogers1, and Eric J. Hunter2

1Department of Communication Sciences and Disorders, University of Utah, Salt Lake City, UT
2Department of Communicative Sciences and Disorders, Michigan State University, East Lansing, MI

In several planned studies, talkers will perform a set of speech tasks several times in a given recording session. For example, one study will have talkers perform the tasks in four conditions (quiet, two levels of noise, and reverberation) in each of two sessions—one in which talkers will be instructed to speak conversationally, and another in which they will be told to speak clearly, as though they were talking to an individual with hearing loss. Speech acoustic characteristics for the various conditions and speaking styles will be compared to determine whether male and female talkers respond differently to changes in the speaking environment or task. Such differences may underlie the notably higher lifetime incidence of voice disorders in women (46% vs. 37% for men).

It is unknown, however, whether or how much the acoustic details of speech are affected by simple repetition of a speech task, or whether such repetition effects differ by gender. Over the course of a recording session, talkers’ speech might become less careful due to fatigue, boredom, familiarity with the speech materials, or some combination of these factors. Such repetition effects could offset the effects of different speaking conditions or of speaking style instructions given at the beginning of a session, and gender differences in repetition effects would need to be accounted for when comparing how men and women accommodate to different environments and tasks.

The present study assessed speech acoustic changes over four repetitions of a speech production task set performed by male and female talkers under conversational or clear speech instructions.

Talkers participated in two recording sessions. During each session, a set of three speaking tasks were performed and repeated four times per test session:

1. The Rainbow Passage
2. A list of 110 sentences
   a) 50 “vowel sentences” (b/v/d/ in neutral context; 5 tokens each of 10 vowels)
   b) 6 lists from the Hearing In Noise Test
3. A picture description task

Although the speaking task order was fixed for the four repetitions, the sentences were in a different random order and a different picture was used each time.

Recordings were made in a quiet, sound-treated room using a headset microphone (Shure SM-10) and a Marantz PMD 670 digital recorder.

In both sessions, talkers were given speaking style instructions (Ferguson, 2004) and a list of 15 practice sentences. Conversational speech was always recorded in the first session, and clear speech in the second session. Talkers were given feedback about conversational speech but not about clear speech.

When talkers felt they were comfortable with the speaking style instructions, they performed the task set a total of four times. They were offered a short break and water to drink after each task set.

For each talker, four measures were taken from the 1st & 4th repetitions of the speaking task set in each style:

1. Vowel space perimeter in Bark: The sum of four Euclidean distances between steady-state F1 & F2 values for the vowels /æ/, /i/, /u/, & /ɪ/ for each task set and vowel, F1 and F2 were extracted from the second, third, and fourth productions of individual vowel sentences using Praat and then averaged.
2. Rainbow Passage Speaking Rate in syllables per second: duration of passage divided by # syllables.

Two pitch measures were estimated from the Rainbow Passage using AudSwiPePrime (Camacho, 2012):

3. Median voice pitch in Hz.
4. Pitch range in Hz: The difference between the 25th and 75th quartiles.

Participants

Talkers were recruited from the University of Utah Department of Psychology participant pool. A total of 19 talkers (9 females) were recorded. A subset of 10 talkers aged 18 to 24 (5 females) were selected for the present analyses. These talkers completed both recording sessions and met the following criteria by self-report:

- They had normal hearing and no history of speech or language disorders.
- They had grown up in Utah and affirmed that “I talk like other people from around here.”

Materials and Procedures

For each task set, four measures were taken from the 1st & 4th repetitions of the speaking task set in each style:

1. Vowel space perimeter in Bark: The sum of four Euclidean distances between steady-state F1 & F2 values for the vowels /æ/, /i/, /u/, & /ɪ/ for each task set and vowel, F1 and F2 were extracted from the second, third, and fourth productions of individual vowel sentences using Praat and then averaged.
2. Rainbow Passage Speaking Rate in syllables per second: duration of passage divided by # syllables.

Two pitch measures were estimated from the Rainbow Passage using AudSwiPePrime (Camacho, 2012):

3. Median voice pitch in Hz.
4. Pitch range in Hz: The difference between the 25th and 75th quartiles.

Results

Gender effects: While the main effect of gender was not significant for vowel space [F(1, 8) = 0.22, p > .769], it just missed significance for speaking rate [F(1, 8) = 4.906, p = .056] and was highly significant for median pitch [F(1, 8) = 196.314, p < .001] and pitch range [F(1, 8) = 20.454, p < .01]. Importantly, gender never interacted significantly with either speaking style or task set.

Vowel space: The main effects of speaking style and task set were significant [F(1, 9) = 9.172 and 7.581 respectively, p < .01]; the interaction between the two effects was not [F(1, 9) = 2.625, p = .14].

Pitch range: The main effect of speaking style was significant [F(1, 9) = 4.66, p < .05]; the task set effect was not [F(1, 9) = 0.38, p = .55]. The interaction just missed significance [F(1, 9) = 4.88, p = .054].

Vowel space: Speaking rate [F(1, 9) = 14.66, p < .01], and task set were significant [F(1, 9) = 2.192 and 2.040, respectively, p > .10].

Acoustic Analyses

For assistance with transcription, acoustic analyses, and data entry: Hannah Jones, Sadie Schlyter, Tyler Jose DaCosta Durães, Jaime Booz, and Miranda Koford at the U of U and Russell Banks, Lauren Glowski, Allison Woodberg, and Amy Kemp at MSU.

For research support: the National Institute On Deafness And Other Communication Disorders of the National Institutes of Health under Award Number R01DC012315. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Acknowledgments

References
