THE DIFFERENT VOICES OF GENDER: SOCIAL RECOGNITION

Nicholas H. Wolfinger
University of California, Los Angeles

Jerome Rabow
University of California, Los Angeles

ABSTRACT

Many researchers have shown that men and women speak differently. In this paper we examine whether these differences extend to the interpretation of speech. Men and women were recorded as they described their participation in a common interpersonal dilemma. From these recordings we transcribed eight short segments of talk that lacked any reference to the gender of the speaker. These quotations were put in a questionnaire administered to 550 UCLA undergraduates. We found that: 1) Respondents were extremely successful at identifying the gender of the speakers; 2) women were better than men at detecting gender; 3) race, socioeconomic status, and other personal characteristics had almost no impact on respondents' ability to identify the gender of the speakers.

INTRODUCTION

The best selling works of Deborah Tannen (1986, 1990; see also 1993a, 1994a, 1994b) have brought renewed interest to a well-established finding: women and men often speak very differently. However, much of the research exploring this gender difference has focused only on speaking. Many researchers have argued that gendered speech should be examined within the context of interaction (see, inter alia, Ochs, 1992; Tannen, 1993b). One dimension of context that has rarely been studied is the interpretation, or hearing, of speech. In this paper we seek to develop an increased understanding of gender differences in communication by asking whether gender impacts hearing as well as speaking. This line of inquiry will contribute to our understanding of speech as an interactive phenomenon.
Many researchers have identified gender differences in speaking styles. For example, Lakoff (1975) notes that women are more likely than men to qualify assertions with tag questions ("It's going to rain tomorrow, isn't it?"). Johnstone (1993) suggests that men's talk is more likely than women's to include references to specific places, times, and objects. Although this sort of finding increases our understanding of gender and communication, it fails to address the interactional nature of conversation. Speaking is only part of conversation; anything said must be interpreted, or heard.

Others have examined gender differences at the level of interaction, but many of these researchers also neglect the issue of interpretation. For example, some researchers argue that men more often interrupt women than vice versa (Zimmerman & West, 1975; West & Zimmerman, 1983; but see James & Clarke, 1993; Tannen, 1994a). Although such findings tell us about the structure of conversation they reveal little about its content; specifically, they do not show how interactants may interpret what many researchers have clearly identified to be gendered speech.

A final impetus for our research comes from the work of Carol Gilligan (1977, 1982), who argues that men and women have distinct voices, or modes of expression. Gilligan claims that women often conceptualize moral dilemmas in terms of social relationships. She uses the metaphor of the web to describe this orientation (Gilligan, 1982, p. 32). In contrast, men typically orient to moral standards or principles, with less reference to social relationships. Gilligan refers to this orientation as a hierarchy of priorities (Gilligan 1982, p. 33).

Gilligan's work has been the source of considerable criticism (for recent reviews see Davis, 1992; Wolfinger, Rabow, & Newcomb, 1993). One response to these criticisms is to question the ways in which researchers, including Gilligan herself, have operationalized the purported gender differences. Employing an essentially psychological paradigm for conceptualizing differences between men and women (Broughton, 1983; Flanagan & Adler, 1983; Johnson, 1983), Gilligan offers limited insight into how the theorized gender differences might emerge in real life interaction; indeed, she herself has acknowledged this shortcoming (1982, p. 126). Like many of the researchers considered here, she does not consider how the gendered voices she identifies might be interpreted in conversation, or heard.

Our research builds on the following premise: if men and women use different voices, they may also be able to hear different voices. By examining listening and comprehension, we hope to learn more about the social context of gendered talk. Can the voices be recognizably heard as well as spoken? Do men and women "hear" differently? If, as Gilligan suggests, women are more attuned to social relationships, they may fare better at voice recognition. Finally, what other factors might explain who hears gendered voices?

To answer these questions we draw on open-ended interviews of men and women discussing their participation in a real life interpersonal dilemma: attempting to prevent another person from driving drunk. From the resulting interview transcripts we selected eight segments of respondent
talk, four from male speakers and four from female speakers. If these transcripts do indeed reflect gendered voices, coders should corroborate the presence of these voices by successfully identifying the gender of the speaker. To test detection of gender we placed the eight segments of transcript in a questionnaire in which we asked respondents to identify the gender reflected by each segment.

This research design has been validated by prior work. Siegler and Siegler (1976) demonstrated that forms of talk identified by Lakoff (1975) as quintessentially female (tag questions) or male (strong unqualified assertions) could often be correctly identified by coders as female or male.[1] Our research expands upon this finding in two ways. First, we have respondents identify, from transcripts, quasi-naturally occurring talk (responses in open-ended interviews) rather than fabricated archetypes. Second, we address the impact of various factors on gender recognition. The most important of these is the gender of the listener: if women and men speak differently they may also hear differently. Additionally we assess the impact of race and social class on voice recognition. Tannen (1986) suggests that people of different ethnicities and social classes may have different conversational styles. More specifically, some critics of Gilligan argue that her research does not take race and social class into account, and therefore applies only to white middle class modes of expression (Nicholson, 1983; Stack, 1986; Tronto, 1987).

This critique could theoretically be extended to most studies of gender differences in speech. Yet these three critics offer no direct evidence in support of their claim. Concrete evidence has been produced by Kollock, Blumstein, and Schwartz (1985), who showed that social power often had a greater impact than gender on the use of tag questions, interruptions, and other linguistic phenomena previously thought to be gender-linked. It follows that race and social class, often markers of social power, may affect talk-in-interaction. We will address these issues by examining whether people of different races and social classes hear differently.

**METHOD**

**Data**

In fall of 1993 we administered questionnaires to volunteers from four undergraduate classes at UCLA. The first section of the questionnaire queried respondents about basic demographic information. Items in this section included age, gender, race, year in school, and social class. We used four items to measure respondents' social class: family income during respondents' senior year of high school, self-assessed social class standing, mother's education and father's education. The second portion of the questionnaire contained the eight segments of transcript that we asked respondents to identify.[2] The eight segments, representing four male speakers and four female speakers, were drawn from twenty-four open-ended, in-depth interviews of peoples' experiences with drunk driving intervention. We selected the eight segments on the basis of four pilot studies, in which we established the likelihood that male and female coders would recognize these eight segments at different rates. The segments varied in length from three to eight lines of typed, transcribed talk. To avoid pressuring respondents we instructed them to
assess each segment on a scale from one to four (1 = definitely male; 2 = probably male; 3 = probably female; 4 = definitely female.) Respondents were not informed about how many of the segments reflected male speakers and how many reflected female speakers.

We obtained 550 questionnaires; 549 of these were usable. Approximately two-thirds of our respondents were female. Although women were clearly overrepresented, the large size of our sample ensured sufficient male respondents for comparative purposes. Our sample was heavily skewed towards more advanced students: only 1% were freshmen, while sophomores, juniors and seniors respectively comprised 20%, 28% and 50.5% of the sample. Graduate students made up the other .5%. The mean age of respondents was 21.

Our sample was quite ethnically diverse: 38% white, 23% Asian American, 19% Latino/a, 7% Black, and 2% Native American. Ten percent selected "Other" and 1% failed to answer. To a lesser degree diversity extended to the social class background of our sample.[3] Six percent of respondents reported their family income to be under 12,000 dollars annually. Thirteen percent identified family income as between 12,000 and 24,000 dollars, while 15% reported $24,000 or more but less than $36,000. Twenty-two percent of respondents fell into the $36,000 to $60,000 range. The modal group of the sample, 24%, fell within the $60,000 to $100,000 range. Finally, 18% of respondents reported an annual family income of 100,000 dollars or more and 2% did not answer this question.

**Analysis**

For the gender identification questions each 4-point answer scale was collapsed into two categories, male and female.[4] Using these categories we coded answers for the eight gender identification items as correct or incorrect. We then tabulated the totals for each item, employing t-tests to show where item totals deviated significantly from the 50% threshold we would have expected from blind guessing. We also computed item totals separately for men and women and used t-tests to verify the presence of significant gender differences. Finally we estimated logistic regression models to examine the effects of various personal and social variables on gender identification.

In all analyses we examine rates of gender recognition separately for each of the eight items. We hypothesized that gendered voices may be more recognizable in some quotations than in others. This information would be lost if we calculated an overall rate of recognition for all eight quotations, an understanding confirmed by the disparate rates of recognition for the different quotations.

**RESULTS**

Are there detectable male and female voices?
Table 1 shows the rates at which respondents successfully identified the gender of the transcript segments. If the transcripts offered no gender clues, then we would expect, given our sample size, identification rates of close to 50% for all eight items. However, our data allow us to reject this null hypothesis: significantly more than 50% of respondents correctly identified the genders of all but one of the quotations. For two of the seven (1 and 6) recognition rates exceeded 75%; for the other five rates ranged from 61% to 68%. That respondents should be able to recognize gender from such short excerpts, transcribed and out of context, suggests that distinct male and female voices exist and can be heard.

Table 1. Success Rates for Gender Coding Task

<table>
<thead>
<tr>
<th>Quotation Number and Gender</th>
<th>Success Rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1F</td>
<td>87</td>
</tr>
<tr>
<td>2F</td>
<td>68</td>
</tr>
<tr>
<td>3M</td>
<td>65</td>
</tr>
<tr>
<td>4F</td>
<td>61</td>
</tr>
<tr>
<td>5M</td>
<td>63</td>
</tr>
<tr>
<td>6M</td>
<td>79</td>
</tr>
<tr>
<td>7F</td>
<td>68</td>
</tr>
<tr>
<td>8M</td>
<td>51</td>
</tr>
</tbody>
</table>

Note. n = 549
All totals significantly (p < .001; 1-tailed tests) higher than 50% except for quotation eight.

Table 2 presents the rates at which men and women successfully identified the gender of the transcript segments. Women were better at identifying the gender for five out of the eight segments, significantly so for two of these five segments (3 and 7), and marginally significantly for a third (4). In contrast, men were more successful than women on only two of the eight segments (6 and 8). Furthermore, for only one of these two did the gender difference even approach significance (quotation 8, p = .09).

Table 2. Success Rates for Gender Coding Task, by Gender
<table>
<thead>
<tr>
<th>Quotation Number and Gender</th>
<th>Female Success Rate (%)</th>
<th>Male Success Rate (%)</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1F</td>
<td>88</td>
<td>85</td>
<td>3</td>
</tr>
<tr>
<td>2F</td>
<td>69</td>
<td>65</td>
<td>4</td>
</tr>
<tr>
<td>3M</td>
<td>70</td>
<td>56</td>
<td>14a</td>
</tr>
<tr>
<td>4F</td>
<td>64</td>
<td>55</td>
<td>9b</td>
</tr>
<tr>
<td>5M</td>
<td>63</td>
<td>63</td>
<td>0</td>
</tr>
<tr>
<td>6M</td>
<td>77</td>
<td>83</td>
<td>-6</td>
</tr>
<tr>
<td>7F</td>
<td>71</td>
<td>61</td>
<td>10c</td>
</tr>
<tr>
<td>8M</td>
<td>48</td>
<td>56</td>
<td>-8d</td>
</tr>
</tbody>
</table>

Note. Total n = 549; 364 women and 185 men
Differences between male and female success rates significant at:
a: p = .001
b: p = .051
c: p = .014
d: p = .093
All tests are 2-tailed.

Table 2 raises a puzzle. Although women seem better at identifying gender from the quotations, why do they do better most but not all of the time? As a partial answer we propose that some of the eight segments may reflect voices that are not strongly associated with gender. Barrie Thorne's (1993, pp. 158, 64) recent conceptualization of gender is useful here:

\[56\]
\[57\]

Those who use a discourse of 'the pinks and the blues,' continually searching for contrastive differences, may assume that girls and boys sharply divide as two separate and unitary types of beings. But the social world is not that simple. ... gender waxes and wanes in the organization of group life, and that flux needs close attention.

We propose to extend Thorne's argument to our research by suggesting that gender may not be always be salient in the different voices represented by the eight segments. Our data support this contention. Table 2 shows that for three of the eight quotes (1, 2, and 5) recognition rates varied little by gender or not at all. Furthermore, Table 1 reveals considerable variation in the overall recognition rates: as much as 36% between quotations (1 and 8). Perhaps gender is more relevant in the quotations with higher recognition rates. However, the explanation proffered here for variability in recognition rates is only speculative: without a more thorough understanding of the substance of the quotes we cannot know for certain why rates of recognition varied so dramatically across the eight segments.
What other factors might influence the recognition of gendered voices? To address this question we used logistic regression to examine the effect of various demographic variables on success rates for each of the eight gender coding items. Table 3 depicts the results of this analysis.[5] Almost none of the five independent variables (age, sex, year in school, social class, or race) predict success for any of the gender recognition items. Social class did predict success for identification of quotation six. However, at 1.07 the odds multiplier (exp[.07] = 1.07) indicates a fairly small effect on the dependent variable and this coefficient is only marginally significant at p < .10. With this exception, the only three significant coefficients are those corresponding to the gender differences we have already described. From these poorly fitting models we conclude that the demographic variables, aside from gender, have a minimal effect on gender recognition. This suggests that hearing the voices does not depend upon, among other things, the listener's race or social class.

Table 3. Logistic Regression of Success Rates for Gender Coding Task on Various Demographic Variables

<table>
<thead>
<tr>
<th>Gender Recognition Items</th>
<th>1F</th>
<th>2F</th>
<th>3M</th>
<th>4F</th>
<th>5M</th>
<th>6M</th>
<th>7F</th>
<th>8M</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>0.06(0.06)</td>
<td>0.01(0.03)</td>
<td>0.00(0.03)</td>
<td>-0.03(0.03)</td>
<td>-0.02(0.03)</td>
<td>-0.01(0.03)</td>
<td>0.00(0.03)</td>
<td>-0.03(0.03)</td>
</tr>
<tr>
<td><strong>Male</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>-0.01(0.28)</td>
<td>-0.21(0.21)</td>
<td>-0.56**(0.20)</td>
<td>-0.40*(0.20)</td>
<td>0.13(0.20)</td>
<td>0.38(0.25)</td>
<td>-0.57**(0.20)</td>
<td>0.18(0.20)</td>
</tr>
<tr>
<td><strong>Freshman or Sophomore</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Freshman or Sophomore</td>
<td>0.28(0.37)</td>
<td>0.42(0.27)</td>
<td>-0.23(0.25)</td>
<td>0.07(0.25)</td>
<td>0.10(0.25)</td>
<td>-0.18(0.29)</td>
<td>-0.19(0.25)</td>
<td>-0.17(0.24)</td>
</tr>
<tr>
<td><strong>Social Class</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social Class</td>
<td>0.03(0.05)</td>
<td>0.03(0.04)</td>
<td>0.04(0.03)</td>
<td>0.01(0.03)</td>
<td>0.00(0.03)</td>
<td>0.07+(0.04)</td>
<td>0.02(0.03)</td>
<td>0.03(0.03)</td>
</tr>
<tr>
<td><strong>Race</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Latino</td>
<td>-0.12(0.52)</td>
<td>0.32(0.40)</td>
<td>0.22(0.38)</td>
<td>0.17(0.37)</td>
<td>0.51(0.38)</td>
<td>0.54(0.48)</td>
<td>-0.03(0.40)</td>
<td>0.04(0.36)</td>
</tr>
<tr>
<td>Black</td>
<td>0.40(0.74)</td>
<td>0.02(0.50)</td>
<td>0.36(0.51)</td>
<td>-0.01(0.48)</td>
<td>0.12(0.48)</td>
<td>-0.33(0.55)</td>
<td>-0.47(0.50)</td>
<td>-0.04(0.46)</td>
</tr>
<tr>
<td>Asian</td>
<td>-0.02(0.49)</td>
<td>-0.04(0.36)</td>
<td>0.06(0.34)</td>
<td>-0.15(0.34)</td>
<td>-0.51(0.34)</td>
<td>-0.40(0.41)</td>
<td>-0.44(0.36)</td>
<td>-0.20(0.33)</td>
</tr>
<tr>
<td>White</td>
<td>0.03(0.47)</td>
<td>-0.40(0.34)</td>
<td>0.03(0.33)</td>
<td>-0.15(0.32)</td>
<td>0.04(0.32)</td>
<td>-0.21(0.40)</td>
<td>-0.36(0.35)</td>
<td>0.25(0.31)</td>
</tr>
<tr>
<td>Native American</td>
<td>0.06(1.15)</td>
<td>-0.28(0.79)</td>
<td>1.40(1.11)</td>
<td>0.07(0.79)</td>
<td>-0.48(0.76)</td>
<td>-0.39(0.90)</td>
<td>-0.55(0.80)</td>
<td>-0.51(0.78)</td>
</tr>
</tbody>
</table>
DISCUSSION

This paper offers four findings: 1) Women and men frequently hear as well as speak different voices; 2) Women seem more sensitive than men to these different voices; 3) Some voices may not be gendered. This would explain why rates of recognition varied so dramatically for the eight quotes; 4) Race and social class do not affect peoples’ ability to hear the voices.

Many researchers find that men and women speak differently. In our paper we show that gender differences in speech can often be heard. We find it noteworthy that these differences pervade speech to the extent that gender is recognizable in short, context-free segments of transcribed talk. Furthermore, we show that the hearing process is sometimes itself gendered. These findings provide new insight into the role of gender in conversation: gender is part of listening as well as talking.

Our results also speak to those who have criticized Gilligan's work on the basis of race and social class (e.g. Nicholson, 1983; Stack, 1986; Tronto, 1987). We found that race and social class had almost no impact on respondents’ ability to recognize the different voices. This finding suggests that some gender differences in communication cut across demographic boundaries.

If women and men speak and hear differently, then they differ in their styles of conversation. This result supports those who claim that men and women typically assume different conversational roles (e.g. Fishman, 1983; Tannen, 1986, 1990, 1994b). Tannen also explores how gender differences in styles of speech can lead to conversational misunderstandings. Such misunderstandings can produce friction between women and men. Greater understanding of exactly how people hear the different voices might provide insight into communication difficulties between the sexes.

Much remains to be learned about the different voices. Perhaps the biggest question concerns the differences between gendered and non-gendered voices. What makes any given quotation recognizable? In the course of conversation, do people use the same cues that make talk observably gendered to linguists? Researchers attempting to answer these questions should be mindful of the fact that talk-in-interaction is collaboratively produced by its participants.
(Goodwin, 1979; Sacks, Schegloff, & Jefferson, 1974). Listening is only one component of conversation. More broadly, researchers studying gender differences in conversation should take the greater social context into account (Ochs, 1992; Tannen, 1993b). To this end, examining isolated segments of interview transcript is only one step in developing a fuller understanding of the relationship between language and gender.

ENDNOTE

[1] In a more recent study coders were unable to identify the authors of short vignettes as male or female (Mulac and Lundell 1994). However, this result is irrelevant to the current study for two reasons. First, Mulac and Lundell's vignettes were prose, and not based on spoken language. Second, the average vignette length was about 120 words, whereas the current study uses much shorter segments of transcript. Longer excerpts are likely to contain numerous and perhaps contradictory gender markers.


[3] Since the four items measuring social class formed a highly reliable scale (see Appendix B), we only present one of them for the purpose of describing the sample.

[4] By collapsing the categories we lose information on the certainty with which respondents identified the gender depicted by the transcripts. This information, however, is a by-product of our research design. We use four-point scales for the answer fields simply to avoid pressuring respondents into failing to attempt identification of the more ambiguous quotations.


APPENDICES

Appendix A. The eight transcript segments. The letter after each number indicates the gender of the speaker.

1F) Q: What did you foresee as the consequences of his driving... [while intoxicated]?
   A: I was worried about him and whoever else would be on the road.

   Q: Was he going to have any other passengers in his car?
   A: No, I do not think so. And it was not a far drive, but he was obviously in no condition to get in the car at all...

2F) I won't go out with a group of people, unless it's already been established... [who will be the designated driver], because, once people have had alcohol, they're not rational beings. So I like to establish what's going on, before we go. As a result, most of my friends have picked up
that habit. So as we're walking out the door, someone says, 'I'm d[esignated] driver,' and, 'so and so has the keys.'

3M) I was a bit concerned. I felt that maybe I ought to nip it in the bud, before it becomes a problem, before he got belligerent. I just sort of told myself, "he's not driving." It was just overall concern. I said to myself, "hey, wake up, let's do something about this."

4F) I just knew that everyone there shouldn't be drinking anymore. It was getting late, and they were going to be going home, though a lot of people spent the night.

5M) Q: What happened between the time you suspected he was drunk and the time that he wanted to leave and you intervened?
   A: Until he voiced his desire to leave, I wasn't paying any attention to him. So I guess nothing really happened. However, when he did get up to leave, I began to wonder if he was O.K. to drive. I asked him, and he said that he was O.K.

6M) Q: What were the consequences you feared of their driving?
   A: Well, with the uncle, it was obvious that, in his condition, he wasn't going to make it very far at all. Hopefully, he wouldn't have even been able to get the key in the ignition. The brother. . . . He wasn't that far away. . . .

   Q: How far is it when you say 'not too far away?'
   A: The brother only had to go about five miles.

7F) So I quietly told my other friends, 'Don't let him drive.' They said, 'He'll be fine.' I said, 'Don't let him drive, I don't want him to.' They talked to him: 'Hey, Joe, why don' you stay here. You can kick back at any one of our places. You probably shouldn't be driving.' He said, 'I'm fine, I do this all the time. I'm great. I'm under control.' I said, 'No.' [after Joe failed to respond to their request]

8M) Q: What was it that you were worried about, if he was drunk?
   A: First of all, I would say: killing himself and any of his passengers, and if you think about it, killing someone else. But the first thing I would think about would be killing himself, or anybody in his car, and the second thing I would think about would be killing anyone else.

Appendix B. Coding of variables for regression.

We treated age as a continuous variable in the regression analyses. Sex was coded as a dummy variable, with female as the reference category. Year, referring to class status at UCLA, was also dummy coded: '1' for freshmen and sophomores, '0' for juniors, seniors, and graduate students. For the race variable we treated the five respondents who failed to answer this question as having
given their race as other. We then coded race as a set of dummy variables, with other as the reference category.

Social class is a scale composed of four measures: self-appraised class status, family income, mother's education and father's education. These were standardized and summed to form the scale, which was highly reliable (Cronbach's Alpha = .83). For the regression analysis we treated each of the eight gender recognition items as a dependent variable in a separate equation. Each gender recognition item, coded 'correct' or 'incorrect', comprised a dichotomous dependent variable; thus, we used logistic regression (see Aldrich and Nelson, 1984).

Finally, we deleted listwise all cases with missing values on any variable. This yielded an n of 494 for the regression analyses.

REFERENCES


AUTHORS' NOTE

Direct correspondence to Nicholas H. Wolfinger, Department of Sociology, 264 Haines Hall, UCLA, 405 Hilgard Avenue, Los Angeles, CA 90095-1551 (e-mail: wolfin@nicco.sscnet.ucla.edu). The authors thank Matthew McKeever, Andrew L. Roth, and two anonymous reviewers for their incisive comments on earlier drafts of this paper. We also express our appreciation to Danielle Edwards for research assistance. A previous version of this paper was presented at the 1994 annual meeting of the American Sociological Association, Los Angeles.

AUTHOR BIOGRAPHIES

Nicholas H. Wolfinger is a doctoral student in the Department of Sociology at UCLA. He is currently finishing a dissertation on the intergenerational transmission of divorce. His research interests include sociology of the family, social psychology, and research methods.

Jerome Rabow is Professor Emeritus of Sociology at UCLA, where he teaches courses in race, psychoanalytic sociology, social psychology, and the sociology of education. He is also a practicing therapist. His research interests lie in race, gender and money, and college students' drinking and driving. He received the Luckman Distinguished Teaching Award from UCLA in 1995 and the Pacific Sociological Association Teaching Award in 1996.