

# Speaker sex and perceived apportionment of talk

**ANNE CUTLER**

*MRC Applied Psychology Unit*

**DONIA R. SCOTT**

*University of Sussex*

## ADDRESS FOR CORRESPONDENCE

Anne Cutler, MRC Applied Psychology Unit, 15 Chaucer Road, Cambridge CB2 2EF, England

## ABSTRACT

It is a widely held belief that women talk more than men; but experimental evidence has suggested that this belief is mistaken. The present study investigated whether listener bias contributes to this mistake. Dialogues were recorded in mixed-sex and single-sex versions, and male and female listeners judged the proportions of talk contributed to the dialogues by each participant. Female contributions to mixed-sex dialogues were rated as greater than male contributions by both male and female listeners. Female contributions were more likely to be overestimated when they were speaking a dialogue part perceived as probably female than when they were speaking a dialogue part perceived as probably male. It is suggested that the misestimates are due to a complex of factors that may involve both perceptual effects such as misjudgment of rates of speech and sociological effects such as attitudes to social roles and perception of power relations.

According to proverbial wisdom, women talk more than men. The English proverb "Women's tongues are like lambs' tails - never still" has parallels in many cultures (Swacker, 1975). But experimental evidence does not justify this belief. Some evidence seems at first to favor it: Preston and Gardner (1967) conducted a factor analysis across a large number of language performance measures and found significant sex differences on a word productivity factor, such that women produced more words than men, paused less often than men, and had a larger vocabulary than men. Similarly, Gall, Hobby, and Craik (1969) found that women produced more words to describe visual displays than men. But in a similar study by Swacker (1975), men produced more words to describe pictures than women, while Brotherton and Penman (1977) found no significant difference between the sexes in a similar task. In each of these investigations, however, the speakers produced monologues, and monologues are not the most frequent type of speech performance. The picture from research on conversations is clearer. Hilpert, Kramer, and Clark (1975) measured the relative contributions of

each partner in a mixed-sex dyad to an unstructured conversation; more often than not, the male partner spoke more. Argyle, Lalljee, and Cook (1968) also found that males in mixed-sex dyads spoke more than females. Markel, Long, and Saine (1976), on the other hand, looked at same- and mixed-sex dyads, and reported no effects for overall proportion of contribution to the conversation; but they did find that female turns were longer. Interestingly, they found that *listener sex* was a significant factor; everyone spoke more to a female listener. By implication, this tells us that there was more inequality in mixed-sex dyads - with males speaking more than females - than in same-sex dyads. In the dyads studied by Duncan and Fiske (1977), it was the men who tended to take longer speaking turns.

In larger groups, the evidence even more clearly suggests that men tend to speak more than women. Eakins and Eakins (1976) found that men spoke longer and more often in faculty meetings than their female colleagues. Spender (1979, 1980) taperecorded spontaneous group discussions and found that male participants often contributed a greater proportion of the discussion even when they were in a minority.

On balance, then, there is actually better evidence for men speaking more than women than vice versa. There is certainly no evidence to support the widespread folk belief that women are overwhelmingly the more garrulous sex.

Perhaps the existence of such a belief itself constitutes evidence that people are very bad at judging how much is spoken. In fact, Hilpert, Kramer, and Clark (1975) found that the perceptions of their conversational participants were in general fairly accurate about who had spoken most, although the men's perceptions were slightly less accurate than the women's. Spender (1979), however, reported very inaccurate perceptions by her group discussants; in a group in which almost twice as many men as women had spoken, participants reported that the majority of speakers had been female.

Extraordinarily, we do not know *how* listeners actually assess how much is spoken. Common sense tells us that someone who drawls a sentence slowly is not considered to have said more than another person who gabbles the same sentence twice in less time than the first person took to say it once. That is, we normally make allowance for speaking rate in judging who says most; amount of linguistic material produced is what really counts.

Although speech scientists have not directly investigated how listeners assess amount of talk, there *is* a large literature on how listeners judge rate of speech. This literature shows that listeners are in fact quite poor at estimating rate. Their estimates take into account both articulation rate (words, syllables, or segments per unit time) and the number and duration of pauses (Grosjean & Lane, 1974, 1976, 1981), but they are also affected by irrelevant factors - either higher intensity or higher fundamental frequency, for example, can cause one of two speech signals spoken at identical rates to be perceived as having a faster rate (Bond & Feldstein, 1982; Feldstein & Bond, 1981), as can absence of intonational phrasing (Rietveld & Gusenhoven, 1987).

The tendency for speech rate to be misjudged as a function of fundamen-

tal frequency is explained as due to our frequent experience of within-speaker concomitant rises in speech rate and voice pitch. But note that we also experience voice pitch variation *between* speakers. In particular, female vocal tracts are generally shorter than male vocal tracts, which makes female voices in general higher pitched than male voices. The tendency for higher pitch to be judged as signaling faster rate suggests that one might find female rates of speech to be systematically judged to be faster than they actually are.

Indeed, Kramer (1977), who studied stereotypes relating to speech, found that not only was "talking a lot" reported to be a characteristic of women, but so was "talking fast." There is no evidence that the second stereotype is any truer than the first; none of the studies reviewed above found any evidence of sex-related differences in the rate of production of words (Brotherton & Penman, 1977; Markel, Long, & Saine, 1976; Swacker, 1975, all explicitly measured this factor). Possibly, then, women are judged to speak faster simply because they have higher pitched voices.

Now what will happen as a result of rate misjudgments if, as we suggested earlier, listeners usually make allowances for rate of speech in judging how much people talk? Obviously, allowances will sometimes be made without justification. One might argue, then, that the folk belief that women talk more than men is just due to rate misjudgment: women do not actually talk more, but they are judged to be talking more because they are judged to be speaking faster, and faster speakers must be talking more per given unit of speaking time.

An alternative suggestion is more complex and may rely on a difference in content between men's and women's speech. Kramer (1975) and Spender (1980) suggested that women are undervalued in society, and as a consequence women's speech is undervalued - female contributions to conversation are overestimated because they are held to have gone on "too long" relative to what female speakers are held to deserve. Preisler (1986) similarly argued that evaluation of women's speech is a function of (under)evaluation of the social roles most usually fulfilled by women.

The former explanation suggests that overestimation of women's conversational contributions is a perceptual bias effect that should be reproducible in the laboratory simply by asking listeners to judge amount of talk produced by male and female speakers, even if content of the talk is controlled. The latter explanation would predict that the content of conversation should make a strong contribution to misjudgments; if content is strictly controlled, there may be no overestimation of women's contributions, but if content is biased, judgments of relative contribution should be similarly biased.

The present study provides an initial test of these predictions. We recorded a number of identical two-party conversations and systematically varied speaker sex. We then had listeners judge the relative proportional contribution of each speaker to the conversation. We could thus ascertain whether male or female speakers were judged as having spoken for a greater proportion of the conversation than was actually the case. If underlying the folk

belief is a perceptual bias that we can reproduce in the laboratory when conversational content is controlled, then we should find that in a conversation in which women and men contribute equally, the women are perceived as talking more than the men.

We also separately estimated the degree to which the textual content of our conversations was judged to be typical of female or of male speech. If the folk belief rests crucially on perception of speech in terms of social roles, then we should find that the degree to which a given contribution is overestimated should be related to how "female" the text of the contribution is judged to be.

It would be desirable to produce speech in which the actual contributions of male and female speakers were exactly matched. This can be done by using synthetic speech, or natural speech that has been digitized and then appropriately compressed or expanded. In either of these cases, however, the speech produced does not sound fully natural. This opens up the possibility that listeners may react to such speech differently from the way they react to speech of "real" men and women speakers. As the present study represents the first direct examination of whether content-controlled speech is misjudged, we preferred to avoid such a possible confounding effect of unnaturalness, and hence we did not digitally manipulate our recorded conversations.

By controlling conversational content, we factored out effects specific to personal participation; our listeners were mere observers of others' interaction, not participants themselves. If they are systematically inaccurate in their perceptions of relative amount of talk, their bias is a general one, not one based on their evaluation of their own participation.

We also tested whether male and female *listeners* varied in the accuracy of their judgments. At least one study (Hilpert, Kramer, & Clark, 1975) suggested that women's perceptions of relative amount of talk could be more accurate than men's. Finally, we varied in addition the sex of the *experimenter*. Our subjects were tested in groups; in such situations the experimenter stands at the front of the room to give instructions and implicitly wields a certain amount of authority. It is possible that the sex of the person giving instructions could affect the operation of any general sex-related bias in conversational perception by affecting transitory perceptions of power relationships. Both authors of this study are female; we therefore enlisted two male colleagues to serve as experimenters for half of the group testing sessions.

## METHOD

### *Subjects*

One hundred thirty-five subjects took part in the experiment, of whom 21 were University of Sussex undergraduates who participated as part of a class project, while the others were volunteers from the subject panel of the Applied Psychology Unit, who were paid a small sum for their participa-

tion. The subjects were tested in groups that varied in size from 10 to 22 members. Each group contained both male and female subjects. Two groups were tested at the University of Sussex and the remaining six at the Applied Psychology Unit. Each of four tapes was heard by two groups, and for each tape, one group was tested by a male experimenter and one by a female experimenter. In all, 60 male and 75 female subjects were tested, and the four tapes were heard by 39, 36, 31, and 29 subjects, respectively.

### **Materials**

We selected four short excerpts from plays, each involving two speakers (in some cases small alterations were made to the original text). Three of the excerpts had roughly the same number of words spoken by each participant; in one case (Dialogue 2), the first speaker contributed approximately twice as much as the second speaker. The four excerpts are reproduced in the Appendix.

Two male and two female actors (members of a university dramatic society) made audiorecordings of the texts. Each dialogue was recorded in four versions:

First speaker female, second speaker male;  
first speaker male, second speaker female;  
both speakers female; and  
both speakers male.

Four tapes were made up from the original recordings. Each tape contained each of the four dialogues and each of the four speaker-order combinations. The order of the dialogues and of the speaker-order combinations was different on each tape. Preceding the dialogues on each tape were two short musical extracts; these were the same on each tape, and served as dummy trials for the experiment.

The real durations of each speaker's contribution to each conversation were measured; the conversations were digitized and stored on disc in a computer, and each turn measured to the nearest centisecond. This demonstrated that our speakers differed considerably in their rates of speech; a given part (the sum of a speakers' turns in a particular dialogue) could take as much as one and a half times as long for one speaker to say as for another. In general, the two females spoke more slowly than the two males (although this was not true of every part).

The texts of the dialogues were also pre-tested for intrinsic textual bias. Fifty-five subjects (25 males, 30 females), none of whom participated in the main experiment, performed this task. Some of these subjects were members of the Applied Psychology Unit subject panel, who were paid for participating; the rest were staff or postgraduate students at the University of Sussex or the Applied Psychology Unit, who were not paid. The subjects were given a transcript of each conversation, with the speakers identified only as *A* and *B*. Below each transcript were the words, "One of the speakers in the above

Table 1. *Text judgment pretest: Proportion of judgments that first speaker was male*

	All subjects	Female subjects	Male subjects
Dialogue 1	.49	.43	.56
Dialogue 2	.15	.20	.08
Dialogue 3	.44	.60	.24
Dialogue 4	.49	.30	.72

dialogue is female, one male. Which is which? Please write A or B in the appropriate box below," and boxes labeled "female" and "male," respectively. On half of the sets of transcripts the words "female" and "male" in the instruction and labels appeared in the reverse order. The four transcripts were stapled together in many different orders.

Table 1 presents the proportions of subjects judging the first speaker to be male for the four dialogues, as well as the same measure for male subjects only and for female subjects only. If the dialogues were not particularly sex-marked, so that either part in each dialogue could be equally plausibly spoken by a male or by a female, we would expect that each dialogue should receive approximately 50% "first speaker is male" judgments in the text task - say between 40% and 60%. Three of our dialogues fell within this range, but Dialogue 2 is biased towards the first speaker being judged as female. Dialogue 2 is in fact the only one in which the actual amount of the two speakers' contributions (measured in number of words) is seriously mismatched; the first speaker has more to say. (However, although it is possible that this factor is in part responsible for the bias, we suspect that a stronger factor is the suggestion of a particular social role, namely, the expectation implied in the final exchange to the effect that the first speaker will be preparing dinner.)

Another potentially important effect appears in the breakdown of the text judgments by subject sex. Although Dialogue 4 received a total of 49% "first speaker is male" judgments, 72% of male subjects thought the first speaker was male, while 70% of female subjects thought the first speaker was female. Similarly, the first speaker of Dialogue 3 was judged to be male by 44% of subjects overall, but by 60% of female subjects and only 24% of male subjects. The differences for these two dialogues are statistically significant: for Dialogue 4,  $\chi^2(1) = 10.57, p < .01$ , and for Dialogue 3,  $\chi^2(1) = 7.45, p < .01$ .

The differences between the dialogues that we have found here, particularly the mismatch in two cases between the responses of male and of female judges, suggest that it would be appropriate to consider the results of our experiment not only as a whole, but also dialogue by dialogue, in order to assess whether listener misjudgments of talk vary as a function of the talk content.

### *Procedure*

For each of the first two trials (the musical excerpts), subjects judged (a) the total duration of the excerpt in minutes/seconds, and (b) the proportion of the excerpt devoted to (i) instrumental music alone, and (ii) instruments and vocalist together. For the four conversations, the subjects judged (a) the total duration, and (b) the proportion of total speaking done by (i) the first speaker, and (ii) the second speaker. The judgments were recorded as percentages. The instructions contained an example judgment: "for instance, if you thought a dialogue took about 50 sec, and the first speaker talked about three times as much as the second, you would write: 50 sec, 75%, 25%."

## RESULTS

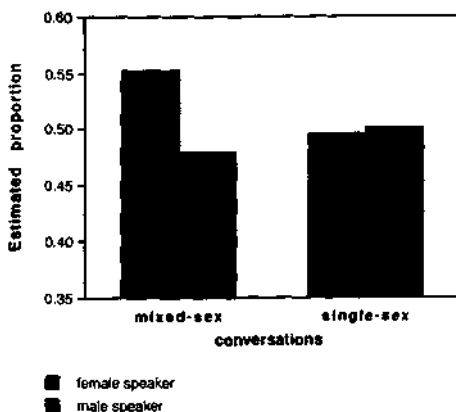
Analysis of apportionment judgments of this kind presents a statistical problem. Subjects in our experiment rated each speaker's proportion of the conversation, but the ratings for each speaker are not independent. For instance, if one speaker in a two-party conversation is rated as taking up 62% of the speaking, it follows that the other speaker receives a rating of 38%. Thus, if we analyze the ratings for both parts of a conversation, we are including in our analysis numbers that are not independent, and this would violate the assumptions of the statistical tests we employ.

The solution to this is, obviously, to conduct our analysis on the ratings assigned to the first speaker or the second speaker only. This does not result in a loss of data, since all of the data excluded can be fully predicted from the data included. We chose arbitrarily to conduct all our statistical analyses on the ratings assigned to the first speaker in each dialogue. Therefore, all of our analyses and the data presented in the figures will refer to the *first speakers only in each dialogue*.

### *Overall results*

We first conducted an overall analysis of the ratings received by the first speakers in all four versions of all four dialogues. The crucial effect here is the interaction between the sex of the speaker and whether the dialogue is single-sex or mixed-sex. If female speakers are perceived as saying more than they actually do, then presumably this effect will cancel out when both speakers are female, but will be apparent when a female is speaking with a male.

Figure 1 shows the mean ratings for male and female first speakers in mixed- versus single-sex conversations. In single-sex conversations, female and male first speakers received almost identical ratings (49.5% and 50%, respectively), but in mixed-sex conversations, female speakers were judged to be talking more (55.2%), male speakers to be talking less (47.8%). Although the number of words spoken was identical for each column, listeners believed that in mixed-sex conversations, females spoke more and males spoke less.



Applied Psycholinguistics 11:3

260

Figure 1. Mean percentage of conversation ascribed to the *first* speaker only, as a function of speaker sex and mixed- versus single-sex participation. In mixed-sex conversations, the female first speaker has a male conversational partner, and the male first speaker has a female conversational partner, while in single-sex conversations, the female first speaker has a female conversational partner, and the male first speaker has a male conversational partner.

In fact, three of these mean ratings are actually underestimates, since the true mean first speaker contribution across all four dialogues was 53.7%. For statistical analysis it is more appropriate to convert the raw ratings to *deviation scores* by subtracting from each rating the true value for the particular dialogue being rated. We conducted separate analyses of variance using subjects and dialogues as random factors. The interaction of speaker sex with whether the dialogue was mixed- or single-sex was significant in both analyses,  $F(1, 119) = 19.32, p < .001$ ;  $F(1, 3) = 51.21, p < .001$ ;  $\min F'(1, 34) = 14.03, p < .001$ . There was also a main effect of speaker sex, with female speakers' contributions being overestimated, but male speakers' contributions being underestimated relative to the actual number of words spoken,  $F(1, 119) = 12.88, p < .001$ ;  $F(1, 3) = 10.6, p < .05$ ;  $\min F'(1, 10) = 5.82, p < .05$ . There was no main effect of subject sex or of experimenter sex, and no further interactions were significant in both analyses.

Separate *t* tests on the components of the interaction revealed that the mean ratings assigned to female and male speakers were significantly different in mixed-sex dialogues,  $t(134) = 5.26, p < .001$ , with subjects as random factor;  $t(3) = 5.89, p < .01$ , with dialogues as random factor; but not significantly different in single-sex dialogues.

Thus, the subjects' ratings were just as predicted by the perceptual bias hypothesis. Although the words spoken by male and female speakers were identical, the female speakers' contributions were overestimated, while the male speakers' contributions were underestimated.



Because the text judgment pretest showed differences between the four dialogues, it is appropriate to assess the effect of content on listener judgments by considering the results separately for each dialogue. The perceptual bias hypothesis predicts an overestimation of female contributions irrespective of content; the social roles hypothesis predicts that overestimations will be greater for parts rated in the pretest as more likely to be female. Because there were no significant differences of any kind in any of the analyses of single-sex conversations, we present the results of these analyses for mixed-sex conversations only.

### **Results for each dialogue**

Figure 2 shows the mean listener ratings of mixed-sex conversations separately for each dialogue. Despite the between-dialogue variations that the pretest revealed, it can be seen that the patterns for all dialogues resemble that for mixed-sex conversations in Figure 1: female contributions were consistently estimated as greater than male contributions.

We performed further analyses of variance, with subjects as random factor, on the deviation scores for each dialogue separately. Several questions are relevant to the consideration of these detailed results:

1. The perceptual bias hypothesis predicts overestimation of female contributions even in parts rated in the pretest (both by male subjects and by female subjects) as equally likely to be spoken by a man or a woman. Dialogue 1 allows this analysis.
2. The social roles hypothesis predicts that if a given part is marked as being female, it is more likely to be overestimated. Dialogue 2 allows this analysis.
3. Where male and female subjects judged a part differently in the pretest, the social roles hypothesis predicts that male and female subjects' judgments in the main experiment should differ similarly. Dialogues 3 and 4 allow this analysis.

*Dialogue 1.* The deviations from actual proportion of words spoken in this dialogue showed that female contributions to mixed-sex dialogues were estimated as greater than male contributions,  $F(1, 59) = 8.77, p < .005$ . No interactions reached significance.

In the pretest, Dialogue 1 revealed *no* intrinsic textual bias. Thus, the significant difference between the ratings for male and female contributions to this dialogue shows that the listener bias can appear with neutral text content.

*Dialogue 2.* In the pretest, both male and female subjects felt that the first speaker's part in Dialogue 2 was likely to be spoken by a female. Thus, the first speaker's part in this dialogue is intrinsically sex-marked as female.

Again, the analysis of deviations from actual proportion of words spoken in this dialogue showed that female contributions to mixed-sex dialogues

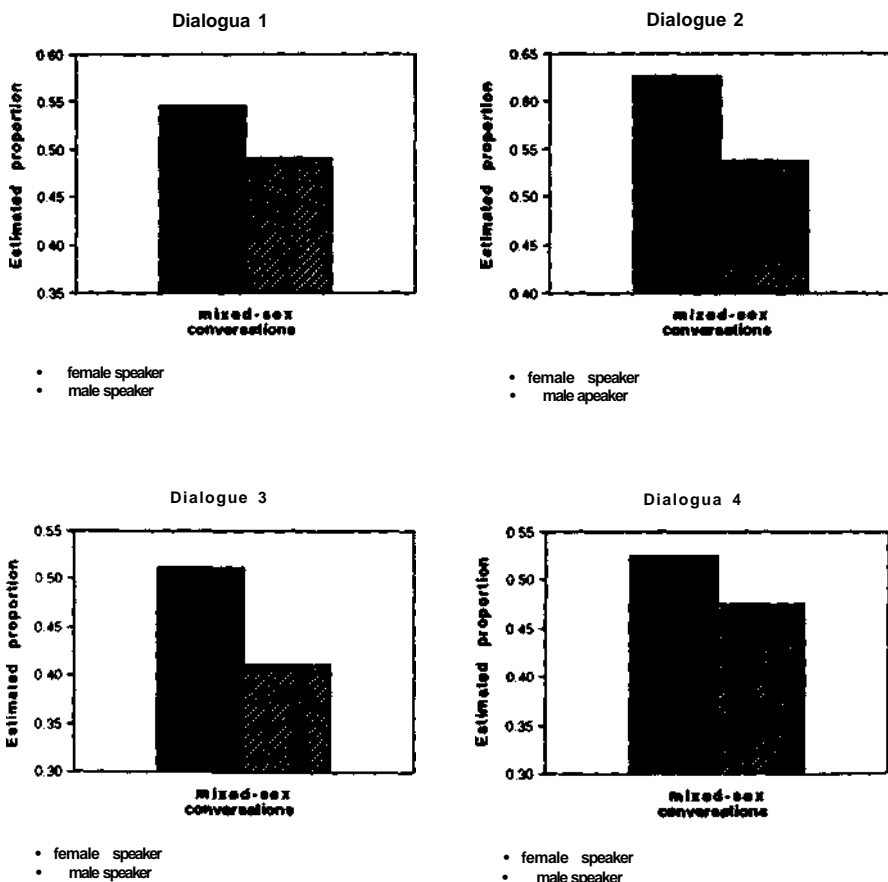


Figure 2. Mean percentage of conversation ascribed to the *first* speaker in mixed-sex conversations, for Dialogues 1 to 4, separately. The female first speaker has a male conversational partner, and the male first speaker has a female conversational partner.

were estimated as greater than male contributions,  $F(1, 59) = 10.69, p < .002$ ; no interactions reached significance.

There was no tendency for the part to be overestimated as a whole; in this dialogue the first speaker's part was actually rather larger than the second speaker's, and all the mean estimates for this part were in fact to a greater or lesser degree underestimates. We tested the size of the male-female ratings difference in this dialogue versus that for Dialogue 1 by analyzing the two dialogues together; there was no significant difference ( $F < 1$ ) in the size of the difference for the two dialogues. On the evidence from this dialogue, therefore, the fact that a particular part appeared to be female-marked seems to have produced no special effects.

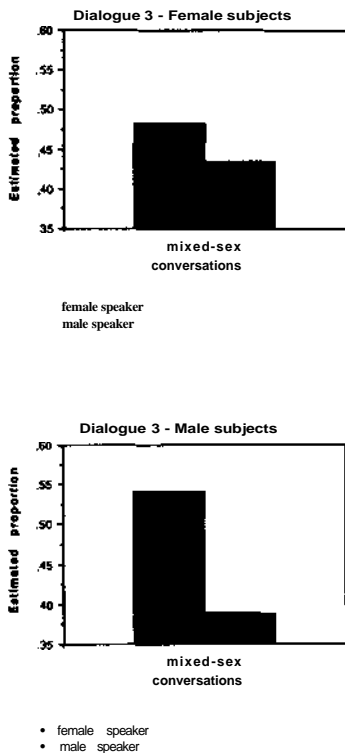


Figure 3. Mean percentage of conversation ascribed to the *first* speaker in mixed-sex conversations, for Dialogue 3, separately for male and female subjects. The female first speaker has a male conversational partner, and the male first speaker has a female conversational partner.

*Dialogue 3.* In Dialogue 3, the pretest judgments showed an asymmetric textual bias. Female subjects thought the first speaker of this dialogue was more likely to be male, while male subjects thought the first speaker was more likely to be female. Thus, it is possible that male and female subjects in the main experiment would also respond differently to this dialogue.

The analysis of deviations from actual proportion of words spoken again found that the difference between the estimates for male and for female speakers in mixed-sex conversations was significant,  $F(1, 60) = 17.19, p < .001$ . As expected, this effect interacted, however, with subject sex,  $F(1, 60) = 4.75, p < .04$ . Further analysis of the components of the interaction showed a significant difference in estimations for male subjects only,  $F(1, 25) = 14.32, p < .001$ ; female subjects showed no such effects. Figure 3 shows the subjects' estimates for this dialogue separately by subject sex.

In contrast with the previous dialogue, the results for Dialogue 3 suggest that the intrinsic textual bias of a part *does* contribute to subjects' estimates. Female subjects did not in this case show the predicted effect; and for female subjects in the pretest, the part appeared to be a male one. This suggests that if a female is speaking a part that is more "male," her contribution is not systematically overestimated. This finding is consistent with the predictions of the social roles hypothesis.

*Dialogue 4.* With this dialogue, there is again a textual bias: males think the first speaker is male, females think the first speaker is female. (The first speaker is scolding the second, who seems, at least early in the extract, to have behaved foolishly.)

The main effect of speaker sex did not reach significance in the analysis of deviations from actual proportion of words spoken,  $F(1, 60) = 3.17, p = .08$ ; the only effect to reach significance was the three-way interaction of subject sex, experimenter sex, and speaker sex,  $F(1, 60) = 4.94, p = .03$ . Within this, female subjects tested by a male experimenter significantly judged female contributions as greater than male contributions,  $F(1, 17) = 5.12, p < .04$ . The other subject groups showed no significant effects; although male subjects tested by a female experimenter also judged female contributions as considerably greater than male contributions, the effect did not reach significance,  $F(1, 13) = 3.28, p < .1$ . Figure 4 shows the subjects' estimates for this dialogue separately by subject sex and experimenter sex.

Although the results for this dialogue suggest that effects of textual bias on perception of a speaker's contribution may be quite complicated, the existence of such effects is in line with the social roles hypothesis.

## DISCUSSION

This study has shown that the erroneous folk belief that women talk more than men is an effect reproducible in the laboratory. Across a large group of subjects and four dialogues, we found that the contribution of a female speaker to a mixed-sex conversation was systematically judged as greater than that of a male speaker, although in fact the contributions were identical. In general, both male and female listeners showed this effect, and they showed it both when they were tested by a male experimenter and when they were tested by a female experimenter.

Earlier, we suggested two possible sources for the mistaken belief. One relied on perceptual bias, the other on interpretation of women's talk in terms of evaluation of women's social roles. The former explanation predicted that female speakers' contributions would be judged to be greater than male speakers' even when they were speaking completely neutral text. Our results showed this to be true and thus support a perceptual bias interpretation. The second explanation predicted that if the content of a part was marked as female, then female speakers' contributions should be particular-

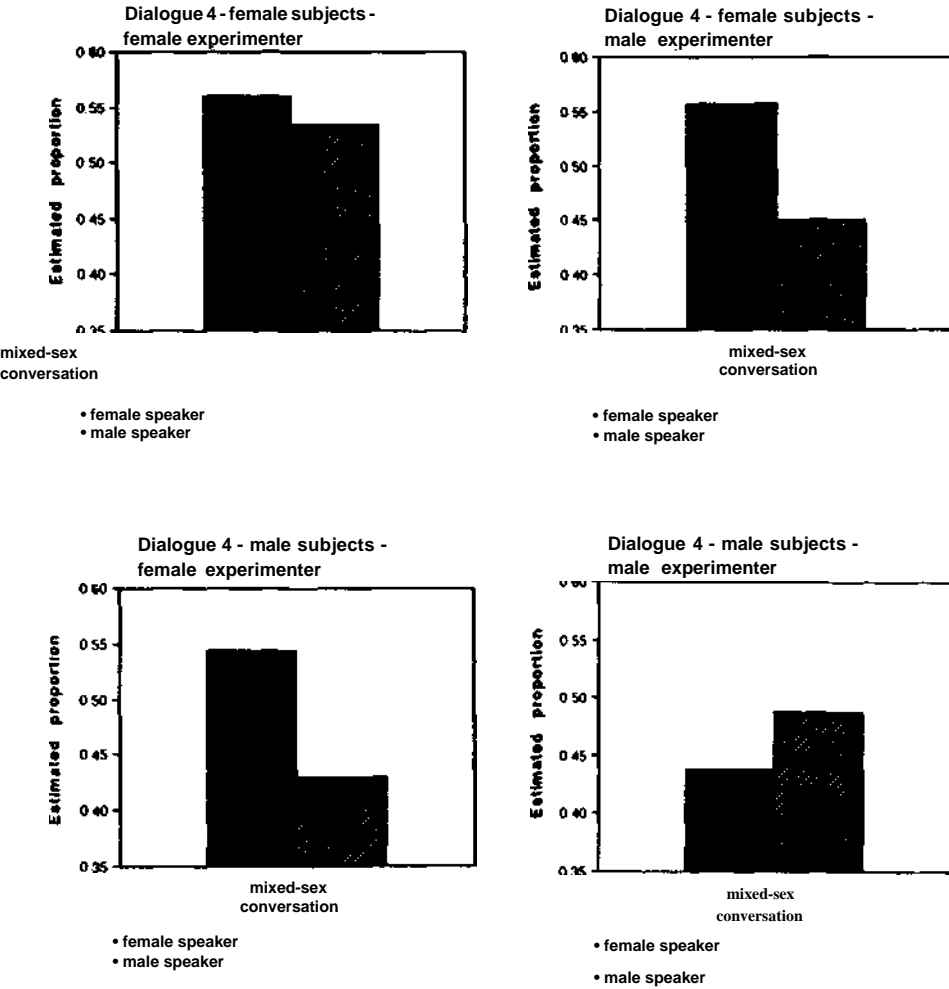


Figure 4. Mean percentage of conversation ascribed to the *first* speaker in mixed-sex conversations, for Dialogue 4, separately for male and female subjects tested by male and female experimenters. The female first speaker has a male conversational partner, and the male first speaker has a female conversational partner.

ly likely to be judged as greater than male speakers' contributions. Our results showed that under certain conditions this was also true, so our results also lend support to a social roles interpretation of why the folk belief exists.

Thus, neither type of explanation has been ruled out by our results. But note also that neither explanation has been directly confirmed. There is

definitely a general bias that can exist independently of content; but we cannot be sure at what level of the perceptual system the bias arises. There is definitely also an interaction of the general bias with the content of talk; but we cannot be sure whether attitudes to women's social roles in general are the source of this interaction. We suggest that one clear conclusion to be drawn from our research is that the genesis of the folk belief that women talk more than men is very complex. Both general classes of explanation that we have proposed are likely to have some truth in them.

Further research is called for to explore the bias we have found and how it interacts with talk content. For instance, in the interests of greater naturalness we did not control speaking rate, voice pitch, or other acoustic parameters in our study. We would hope that our finding of a perceptual bias effect with natural (albeit content-controlled) speech will stimulate others to follow up our work with systematic manipulation of such parameters. Similarly, the social roles hypothesis must be properly tested by systematic manipulation of the content of the speech to be judged, along dimensions known to be related to social roles.

We believe that there are several other aspects of our results that have important implications for research on how talk is perceived. First, there is evidence in our results that our subjects were *not* basing their estimates on judgments of elapsed time. Second, the pattern of content effects we found suggests that estimates for male and for female talk are not symmetric. We will discuss each of these conclusions separately.

#### *Perception of relative amount of talk*

Recall that we measured the durations of each contribution by each speaker in our materials. This allowed us to assess whether our subjects' judgments were related to a contribution's duration.

If this were the case, we should expect the judgments to be accurate when they are considered relative to speaking time; for instance, the grand mean deviation of judged proportion of conversation from actual proportion of speaking time should be close to zero. In fact, it is  $-2.12$ , which is significantly different from zero,  $t(119) = -5.12$ ,  $p < .001$ . Although our results show that subjects' judgments considered relative to words spoken were not at all accurate, they turned out to be even less accurate considered relative to speaking time; the mean unsigned deviation from proportion of words spoken was 9.26 percentage points, whereas the mean unsigned deviation from proportion of speaking time was 10.42 (but these two means are not significantly different). We also correlated subjects' estimates with the actual speaking time measurements, and found that the correlation was by no means high:  $.375$  (although, given the large  $N$ , this figure is significantly different from zero).

All this strongly suggests that our subjects were not basing their judgments of amount of talk on elapsed time. Of course, it may seem somewhat unlikely to propose that subjects should judge speaking time accurately. While studies of human time perception have shown that listeners are, at

least to some degree, able to judge elapsed time in a linear manner (Allan, 1979), time judgments are also known to be governed by psychophysical functions, and certainly this is true of judgments involving the timing of speech (see, e.g., Bond & Feldstein, 1982; Grosjean & Lane, 1976).

Nevertheless, we were concerned to make very sure of this point in order to control for a possible confounding of rate of speech with speaker sex in our study. Although none of our speakers spoke noticeably slowly or noticeably rapidly, and although the effects differed from dialogue to dialogue and within dialogue versions, the durational measurements did show speech rate differences between our female and male speakers. These differences were in the opposite direction to those predicted by the stereotype that women "talk fast": on average, our two female speakers tended to speak more slowly than our two male speakers, although an analysis of variance of the four speakers' rate of speech (words per second) across the eight parts that each played showed that the mean rates for males (3.96 wps) and females (3.26 wps) were not significantly different,  $F(1, 2) = 2.29$ .

If our subjects had attempted to judge amount of talk as amount of elapsed time, therefore (and had succeeded in doing it accurately), they might simply have judged female speakers' contributions as greater because female speakers spoke more slowly and took up more time. This suggestion is in conflict with the apparent bias to perceive women as talking fast and, indeed, in conflict with the everyday experience of making allowance for speaking rate in judging who says most. Nevertheless, we wished to rule out the suggestion definitively.

This could be done by using evidence from one of the four dialogues. Although in general the speaking time differences were in the same direction as the estimates predicted by the experimental hypothesis, this was not the case for all dialogues. In Dialogue 3, in which there is least variation in speaking time across the four versions of the dialogue, all four first speakers spoke more rapidly than second speakers, even the male-partnered female (who took up 47% of the words, but 46% of the speaking time). Thus, there is no way that variation in speaking time could account for the bias effect should we find it in this dialogue.

We therefore analyzed the results of this dialogue as deviations from proportion of speaking time. The effects were just as in the earlier analysis of deviations from proportions of words spoken: female contributions were estimated as greater than male contributions only in mixed-sex conversations and only by male subjects,  $F(1, 25) = 3.91, p = .059$ .

Thus, when speaking time differences could not possibly account for the effect of misestimates of proportion of words spoken, the systematic misestimates were still there; even considered relative to speaking time, subjects' estimates were biased in the same systematic way. Therefore, we feel justified in rejecting the possibility that our subjects' estimates of proportion of talk were based on veridical estimation of elapsed time. Note that this provides at least a preliminary experimental confirmation of the commonsense view of how listeners assess amount of talk: simple estimation of time taken up is not the whole story.

### ***Asymmetry in speaker sex effects***

When a part was not particularly sex-marked (Dialogue 1), females speaking it were judged to have said more than males speaking it. When a part was marked as female for male and for female subjects alike (Dialogue 2), the same effect was found. When, however, a part was marked as female for male subjects only (Dialogue 3), only male subjects showed the effect; and when a part was marked as female for female subjects only (Dialogue 4), only female subjects showed any effect.

What can we learn from this pattern of results? First, consider the fact that the listener bias to judge female contributions as greater than male contributions is no greater for female-marked content (Dialogue 2) than for neutral content (Dialogue 1), nor is a female-marked part overestimated in general, irrespective of who speaks it. This suggests that the effects of content do not operate symmetrically and across the board - a female-marked part is only effectively female-marked if it is spoken by a female. This is in fact not incompatible with the social roles hypothesis, if one assumes that a typically female social role may lose its female-marking if it is performed by a male. Note that we ascribed the female-marking of the first speaker's part in Dialogue 2 for all subjects to the suggestion that the first speaker would be preparing dinner; it seems reasonable to suggest that dinner preparations by a male may not be an undervalued role.

Second, consider the fact that although the listener bias to judge female contributions as greater than male contributions appears with neutral content (Dialogue 1) and with female-marked content (Dialogues 2, 3, and 4), it disappears with male-marked content (Dialogues 3 and 4). On the one hand, this suggests a further asymmetry, namely, that the bias consists of overestimation of female contributions rather than *underestimation* of male contributions. The reason for this is that if male contributions were being underestimated, they would presumably be underestimated in male-marked parts; but in fact we found no significant differences in the listener judgments for male-marked parts.

On the other hand, the absence of effect in male-marked parts implies that typically male content can remove the overestimation of female contributions. This is again consistent with the social roles hypothesis in that it suggests that placing a female speaker in a male (and, by implication, more highly valued) role can remove the effect by which the amount of talk she produces is overestimated. A purely acoustic explanation of the listener bias that we have observed would predict that the effect would remain in male-marked parts. It does not; the effect is asymmetric and, as such, consistent with an explanation in terms of social roles.

Related to this point is the one aspect of our results that remains to be discussed, namely, our finding with Dialogue 4 of an effect of the sex of the experimenter. In this dialogue, the first speaker's part was marked as female for female subjects only. Indeed, only female subjects showed the predicted significant tendency to overestimate female contributions - but then *only when they were tested by a male experimenter*. Looked at from the other



way, female subjects did *not* make misjudgments as a function of speaker sex when they were tested by a female experimenter. It is an intriguing possibility that the presence of female authority, in the person of the experimenter, succeeded in removing, for female subjects, a general devaluation of women's contribution underlying the biasing effect.

We have no independent evidence for this suggestion that perception of female authority can exercise such effects on estimation of amount of talk; and the suggestion raises a good many further questions. Why, for instance, did we find that female subjects with female experimenters *did* overestimate female contributions to other dialogues? Why did male subjects not alter their judgments as a function of the sex of the experimenter? A partial answer might lie in the fact that the text judgment pretest results showed that the female marking of the first Dialogue 4 part for female subjects was slightly less strong than the female marking of the first speaker's parts in Dialogues 2 and 3 - perhaps the effects of female authority as manifested in a female experimenter are not very powerful and can only override content effects when these are weak. This is clearly an issue that can only be settled by means of further research.

Thus, the conclusions to our research must remain somewhat tentative. We have established that the folk belief that women talk more than men can be replicated in the laboratory: listeners show a general bias, independent of content, towards overestimating female speakers' amount of talk. In that this effect appears when content is neutral, it could well have its basis in some perceptual effect such as misjudgment of rate of speech. However, we have also shown that talk content can interact with this bias, in that the bias is more likely to appear with typically female talk than with typically male talk. This is consistent with an explanation of the bias in terms of attitudes to women's social roles. Our data do not permit a definitive choice between these two types of explanation, and we suspect that both may contribute to our findings. Clearly, there is still a good deal of interest to be found out about why women are often perceived as talking more than men, even though they don't.

## APPENDIX

### DIALOGUE 1

A: Now then -

B: Now then what?

A: (contemptuously) Now then what!

B: I don't know what you're talking about.

A: Oh, you don't, don't you?

B: No, I don't, so shut up.

A: I suppose you don't know you mucked up the whole exit.

B: It wasn't my fault.

A: Whose fault was it then, Mussolini's?

B: (with sarcasm) Very funny.

- A: (witheringly) I don't suppose you dropped your prop, did you? And having dropped it, you didn't have to go back for it, leaving me to prance off by meself - who do you think you are, eh?
- B: The exit was too quick.
- A: It was the same as it's always been.
- B: It was too quick. It's been too quick the whole week, the whole number's too quick -
- A: Bert Bentley takes that number at the same tempo as he's always done.
- B: You and your Bert Bentley, just because he stands you a Welsh Rarebit at the Queen's you think he's God Almighty.
- A: Listen, Bert Bentley's the best conductor in the North of England, and don't you make any mistake about it.
- B: Best conductor my foot! I suppose he thinks it's funny to see us leaping up and down the stage like a couple of greyhounds!

### DIALOGUE 2

- A: You mean you've asked him to come over?
- B: Yes, I've asked him to dinner.
- A: You really did?
- B: I did!
- A: You did - and did he - accept?
- B: He did!
- A: Well, well - well, well. That's fine!
- B: I thought you'd be pleased.
- A: It's definite, then?
- B: Very definite.
- A: Soon?
- B: Very soon.
- A: For heaven's sake, stop carrying on and tell me *some* things, will you?
- B: What things do you want me to tell you?
- A: Naturally I would like to know when he's coming!
- B: He's coming tomorrow.
- A: Tomorrow?
- B: Yes. Tomorrow.
- A: But - tomorrow gives me no time!
- B: Time for what?
- A: Preparations! Why didn't you phone me at once, as soon as you asked him, the minute he accepted?

### DIALOGUE 3

- A: I spoke to Robert this morning.
- B: Oh?
- A: I'm taking him to lunch on Thursday.
- B: Thursday? Why?
- A: Well, it's my turn.
- B: No, I meant why are you taking him to lunch?

## Cutler &amp; Scott: Speaker sex and perception of talk

- A: Because it's my turn. Last time he took me to lunch.  
 B: You know what I mean.  
 A: No. What?  
 B: What is the subject or point of your lunch?  
 A: No subject or point. We've just been doing it for years. His turn, followed by my turn.  
 B: You've misunderstood me.  
 A: Have I? How?  
 B: Well, quite simply, you often do meet, or have lunch, to discuss a particular writer or a particular book, don't you? So to those meetings, or lunches, there is a point or subject.  
 A: Well, there isn't to this one.

**DIALOGUE 4**

- A: The classroom is for instruction, not for propaganda. I think that's something you might bear in mind yourself.  
 B: Sorry. I don't follow you.  
 A: The classroom is for instruction. Not propaganda.  
 B: Of course.  
 A: It's something you might well bear in mind yourself.  
 B: I'm sorry. I'm afraid I don't know what you're referring to.  
 A: Perhaps you'd just better think about it for a while. Now if we could -  
 B: I'm sorry, what exactly were you referring to?  
 A: Did you take your students on a tour of Ford's recently?  
 B: Yes.  
 A: Did you follow up that tour with a lecture to the students on the evils of the car industry?  
 B: No, I didn't. Who told you that?  
 A: Look, I respect your point of view, but you're not employed here to ram it down students' throats.  
 B: I'm afraid I take very strong exception *to* that statement! I don't know who was the source of your information, but if they'd reported me accurately - look, I merely stated the fairly obvious truth that the main criteria determining car design are ostentation, model proliferation and manufacturing economy! If that's forcing politics down my student's throats I offer you my resignation right here and now!  
 A: Don't be ridiculous.  
 B: I mean it!

**ACKNOWLEDGMENTS**

Donia R. Scott acknowledges financial support from a grant by the Science and Engineering Research Council, U.K., to the University of Sussex. We thank Dennis Norris and Mark Pearson for serving as additional experimenters, and Mark Pearson also for collaboration in the design and recording of the experimental materials. The reporting of this work has benefited greatly from useful comments by Deborah Cameron, Dennis Norris, Valentina d'Urso, and Peter van Sommers, as well as two

anonymous reviewers for this journal; particularly valuable suggestions were made by David Good, Francois Grosjean, and Virginia Valian. We further thank Ian Nimmo-Smith for statistical advice. D.R.S. is now with Philips Research Laboratories, Redhill, Surrey, U.K.

#### REFERENCES

- Allan, L. G. (1979). The perception of time. *Perception & Psychophysics*, 26, 340-354.
- Argyle, M., Lalljee, M., & Cook, M. (1968). The effects of visibility on interaction in a dyad. *Human Relations*, 21, 3-17.
- Bond, R. N., & Feldstein, S. (1982). Acoustical correlates of the perception of speech rate: An experimental investigation. *Journal of Psycholinguistic Research*, 11, 539-557.
- Brotherton, P. L., & Penman, R. A. (1977). A comparison of some characteristics of male and female speech. *Journal of Social Psychology*, 103, 161-162.
- Duncan, S., & Fiske, D. W. (1977). *Face-to-face interaction*. Hillsdale, NJ: Erlbaum.
- Eakins, B., & Eakins, G. (1976). Verbal turn-taking and exchanges in faculty dialogue. In B. L. Dubois & I. Crouch (Eds.), *The sociology of the languages of American women*. San Antonio, TX: Trinity University.
- Feldstein, S., & Bond, R. N. (1981). Perception of speech rate as a function of vocal intensity and frequency. *Language & Speech*, 24, 387-394.
- Gall, M. D., Hobby, A. K., & Craik, K. H. (1969). Non-linguistic factors in oral language productivity. *Perceptual and Motor Skills*, 29, 871-874.
- Grosjean, R., & Lane, H. (1974). Effects of two temporal variables on the listener's perception of reading rate. *Journal of Experimental Psychology*, 102, 893-896.
- (1976). How the listener integrates the components of speaking rate. *Journal of Experimental Psychology: Human Perception & Performance*, 2, 538-543.
- (1981). Temporal variables in the perception and production of spoken and sign languages. In P. Eimas & J. L. Miller (Eds.), *Perspectives on the study of speech*. Hillsdale, NJ: Erlbaum.
- Hilpert, F., Kramer, C., & Clark, R. A. (1975). Participants' perceptions of self and partner in mixed sex dyads. *Central States Speech Journal*, 26, 52-56.
- Kramer, C. (1975). Women's speech: Separate but unequal? In B. Thorne & N. Henley (Eds.), *Language and sex: Difference and dominance*. Rowley, MA: Newbury House.
- (1977). Perceptions of female and male speech. *Language and Speech*, 20, 151-161.
- Markel, N. N., Long, J. F., & Saine, T. J. (1976). Sex effects in conversational interaction: Another look at male dominance. *Human Communication Research*, 2, 356-364.
- Preisler, B. (1986). *Linguistic sex roles in conversation*. Berlin: de Gruyter.
- Preston, J. M., & Gardner, R. C. (1967). Dimensions of oral and written language fluency. *Journal of Verbal Learning and Verbal Behavior*, 6, 936-945.
- Rietveld, A. C. M., & Gussenhoven, C. (1987). Perceived speech rate and intonation. *Journal of Phonetics*, 15, 273-285.
- Spender, D. (1979, February). Language and sex differences. *Osnabrucker Beitrage zur Sprachtheorie*, 38-59.
- (1980). *Man-made language*. London: Routledge & Kegan Paul.
- Swacker, M. (1975). The sex of the speaker as a sociolinguistic variable. In B. Thorne & N. Henley (Eds.), *Language and sex: Difference and dominance*. Rowley, MA: Newbury House.