BIOLOGICAL GENDER AND SOCIAL GENDER IN RELATION TO VOICE QUALITY

Monique Biemans and Renée van Bezooijen
Nijmegen University, the Netherlands

ABSTRACT

This paper addresses part of a study into the relation between gender and voice quality. Results are presented of perceptual voice quality descriptions by phonetically naive listeners and the attribution of gender-related personality traits on the basis of voice samples, in relation to the speakers' biological gender (the sex of the speakers) and social gender (gender identity). Strong relations were found between the voice descriptions and the gender attributions, especially between 'dynamic' gender-related personality traits (e.g. *enterprising*, *self-assured*) and e.g. the voice features *fast*, *relaxed*, *varied*. Relations between biological and social gender on the one hand and the voice descriptions and the gender attributions on the other hand are marginal.

1. INTRODUCTION

The relation between voice quality and gender has been studied phonetically in several 'production' studies: variation in voice quality between the sexes is described on the basis of acoustic measurements or auditory descriptions by expert listeners. Most research has concentrated on differences in pitch, but also phonation characteristics such as differences in breathiness and creakiness have received attention, and in a more limited manner differences in loudness and tempo have been inventorised [e.g. 5, 6, 9, 12].

Within social psychology the 'perception' side of voice quality has been studied through 'attribution' studies, in which personality traits are assigned to speakers on the basis of their voices. Predominantly the relation between personality judgments and pitch has been looked at, and to a lesser extent the relation with breathiness, creakiness and loudness [e.g. 3, 7, 8, 10]. Not much research has explicitly studied the relation between the attribution of gender-related personal characteristics and voice. An exception is Van Bezooijen [11], who found differences in the evaluation of Dutch and Japanese female speakers by Dutch and Japanese listeners as a function of mean pitch.

In the present study a distinction is made between biological and social gender. Biological gender comprises the anatomical and physiological features which indicate maleness and femaleness, whereas social gender is the construction of masculinity and femininity, i.e. the social, cultural, and psychological characteristics a person acquires by becoming a man or a woman in a specific culture at a specific moment in time.

In phonetics this difference has been acknowledged for some time. On the one hand there is an anatomical cause for e.g. pitch and formant frequency differences between men and women, due to the differences, on average, in the size of the larynx and the vocal tract. On the other hand, the social environment is also known to influence speech behavior. The voices of boys and girls differ in pitch and formant frequencies before they develop sex-specifically in puberty [4]. Furthermore, pitch, and voice quality in general, appear to be culturally variable. Social notions about e.g. gender influence the way people speak (e.g. [11]).

2. COMMUNICATION MODEL

The present research distinguishes itself from earlier gender and voice quality research in three ways: 1. Both the production and perception point of view are incorporated. 2. Multiple voice quality characteristics are investigated simultaneously, instead of only one or two. 3. The combined effects of biological gender (the sex of the speaker) and social gender (gender identity) are examined.

I	II	III	IV	V
the sex of the speaker and gender identity		voice quality characteristics		attributions of gender identity
self-assessment	*acoustic descript.*	*percept. expert descript.*	*percept. naive descript.*	*percept. naive judgm.*

Figure 1. Communication model.

The research has been designed according to the social-psychological model for interpersonal communication proposed by Scherer [8]. This model, adapted for the present study to the communication of gender, can be found in Figure 1. It visualises a continuum from speaker (on the left) to listener (on the right). The five levels represent:

I. The speaker characteristics biological gender (the sex of the speaker) and social gender (gender identity). Hypothetically, these are expressed in the voice of a speaker.

II. Objective acoustic voice quality descriptions: maximum correspondence with the speech signal, minimal certainty that is described what a listener hears.

III. Perceptual voice quality descriptions by expert listeners: valid description on the basis of objective categories, but also influence of subjective listener judgments.

IV. Perceptual voice quality descriptions by phonetically naive listeners: low expected correspondence with objective categories, relatively large certainty that is being described what 'common' people hear.

V. Perceptual gender identity judgments by naive listeners: assignment of gender-related personality characteristics on the basis of a speaker's voice.

The purpose of the research is to gain insight in the relations between all five levels, i.e. the communication of gender through voice quality. In the present article, the relations between levels I, IV, and V are described, i.e. the relations between the gender characteristics of the speakers (the sex of the speakers and their self-assessed gender identity), the perceptual voice quality descriptions by naive listeners, and gender identity attributions by naive listeners.

3. METHOD
3.1. Speech material and speakers
The total speech corpus consists of recordings of 114 speakers divided over 57 interactions, 19 between two women, 19 between two men, and 19 between a woman and a man (Sony TCD-D7 DATrecorder, two Sennheiser microphones MD421N / MD441N). The speakers in each dyad were acquaintances, e.g. friends, room-mates, fellow students, but not partners or family. They spoke for half an hour about designated topics, such as ethical problems (e.g. who should receive a donor kidney) or statements (e.g. smoking in restaurants should be prohibited).

For each speaker a one-minute speech sample was selected. First, a computer program segmented the speech into short utterances.[1] Segmentation criterion was a silent pause of at least 200 ms. Speech files which contained sound from both microphones (simultaneous speech or noise) were discarded, as well as very short (< 1 sec) and very long (> 5 sec) files. Subsequently the speech files with the most neutral content with regard to gender were selected. Content neutrality was assessed in a paper-and-pencil experiment.

For levels IV and V of the research model shown in Figure 1, 48 speakers were selected. Their ages range between 19 and 40, they have a higher professional or university education, are non-smokers, and do not speak with a noticeable regional accent (according to two expert listeners).

3.2. Self-assessment of gender identity (level I)
Social gender was operationalised as gender identity, i.e. the extent to which gender-stereotypical behavior, thoughts or feelings are considered to be part of a person's identity. All speakers filled out a gender identity questionnaire, developed specifically for the Netherlands in the early nineties [13]. The questionnaire has 30 questions about personality traits and 28 about behavior. Answers are given on 5-point scales. The questionnaire assesses stereotypical feminine traits, masculine traits, feminine behavior and masculine behavior.

7-point scales (1-7)	
smooth -	rough
precise -	sloppy
tense -	relaxed
slow -	fast
monotonous -	varied
low -	high (for a woman/man)
4-point scales (0-3)	
not whispery -	whispery
not creaky -	creaky
not nasal -	nasal (through the nose)

Table 1. Voice quality scales.

3.3. Perceptual voice quality descriptions by naive listeners (level IV)
28 naive listeners carried out the perceptual description of voice quality, 14 women and 14 men, aged between 19 and 26. They were students at Nijmegen University, from different disciplines.

They judged the speakers' voices on nine scales (Table 1). The scales were selected from the literature on the basis of their relevance to gender. The terminology was adapted for use by naive listeners. The scale low/high was scored sex-specifically, to prevent that the upper half of the scale would be used for female speakers and the bottom half for male speakers.

3.4. Gender identity attributions by naive listeners (level V)
28 subjects (different from those who made the voice quality descriptions) participated in the attribution experiment, 14 women and 14 men, whose ages ranged between 19 and 28. They were students at Nijmegen University, from different disciplines.

On the basis of the one-minute voice samples, they judged the personality of the speakers on twelve scales (Table 2). The scales were selected from the gender identity questionnaire (used in level I) on the basis of high item-total correlations (central to the concepts masculinity and femininity as defined in the questionnaire) and their distribution over the different factors of a factor analysis (semantic diversity). The scales *masculinity* and *femininity* were added. All scales were 5-point scales, which ranged from 1 = not to 5 = very.

Feminine scales	Masculine scales
emotional	self-assured
warm-hearted	enterprising
sensitive	adventurous
romantic	technical
caring	independent
feminine	masculine

Table 2. Gender identity traits.

4. RESULTS AND DISCUSSION
4.1. Reliability
The reliability of the descriptions and judgments was assessed by means of Cronbach's α. For the voice quality scores it varied from .74 (*nasal*) to .96 (*slow/fast*), which is rather high to very high. For the gender identity attributions reliability ranged from .72 (*warm-hearted*) to .96 (*masculine*), which again is rather high to very high. All scales were scored sufficiently reliably to be included in the rest of the analyses.

4.2. Clustering of scales
4.2.1. Voice quality. The voice quality scales were factor analysed to discover a possible clustering of scales (principal components analysis with varimax rotation) (Table 3. Factor loadings > .6 and < .4 on other factors are in bold, loadings < .1 have been left out.). Two factors were extracted with eigenvalues > 1, together accounting for 63% of the variance. The first is a combination of *slow*, *relaxed*, *low*, and *monotonous* (or *fast*, *tense*, *high*, and *varied*), and is predominantly prosodic, combined with the feature *tension*. The second consists of *smooth*, *precise*, *not creaky*, *not whispery* (or *rough*, *sloppy*, *creaky*, *whispery*), and is predominantly phonational, combined with the feature *precision*.

	Factors	
	1	2
slow/fast	**.86**	-.12
tense/relaxed	**-.81**	-.16
low/high	**.79**	-.32
monotonous/varied	**.72**	-.36
nasal	-.47	.45
smooth/rough	-.29	**.84**
precise/sloppy		**.83**
creaky		**.71**
whispery	-.17	**.65**

Table 3. Factor analysis of the voice quality descriptions.

4.2.2. Gender attributions. The gender attributions were also factor analysed (Table 4. Factor loadings > .6 and < .4 on other factors are in bold, loadings < .1 have been left out.). Three factors were extracted with eigenvalues > 1, together accounting for 86% of the variance. The first is a combination of the feminine traits *warm-hearted*, *romantic*, *caring*, and *sensitive*, and can be summarised by the concept 'emotional warmth'. The second consists of the masculine traits *enterprising*, *self-assured*, *adventurous*, and *independent*, and can be labeled 'dynamic'. The third factor is made up of the two general concepts *masculinity* and *femininity*, and will be called 'feminine', because of its positive factor loading for feminine.

	Factors		
	1	2	3
warm-hearted	**.91**	.17	
romantic	**.83**		.32
caring	**.81**	-.22	.11
sensitive	**.81**	-.37	.35
emotional	.74	-.27	.48
enterprising		**.96**	
self-assured	-.17	**.89**	-.24
adventurous		**.89**	.16
independent	-.35	**.88**	.15
masculine	-.16		**-.95**
feminine	.23		**.94**
technical	-.47	.30	-.65

Table 4. Factor analysis of the gender attributions.

4.3. Relations between levels I, IV, and V
4.3.1. Voice quality and gender attributions (IV and V). Our hypothesis is that voice quality influences gender attributions. Therefore, three stepwise linear multiple regression analyses were performed, with the three factors of the gender attributions (emotional warmth, dynamic, feminine) as criterion variables and the nine voice quality scales as predictor variables. The results are shown in Table 5. The attribution of 'dynamic' gender traits is fairly strongly related to a combination of several prosodic and phonational voice quality features; the model explains 75% of the variance in the factor scores, which is rather high. The model for the factor 'feminine' has a considerably lower percentage of explained variance, 21%, but still one fifth of the variance can be accounted for by the combined scores for rate and whisperiness. The scores for the factor 'emotional warmth' could not be predicted from the voice quality judgments.

The features *fast* and *not whispery* appear both in the model for 'dynamic' and that for 'feminine', which at first sight might seem remarkable. However, the scores for the traits *feminine* and *masculine* were highly correlated with the sex of the speaker, to the point that male speakers only received scores between 3 and 5 on *masculine* and between 1 and 3 on *feminine*, with the opposite pattern emerging for female speakers. (see also section 4.3.3.). Therefore, the factor 'feminine' can in a sense be seen as the equivalent of the sex of the speaker.

Factor	Model	R^2
1. emotional warmth	—	—
2. dynamic	fast	.75
	relaxed	
	precise	
	rough	
	varied	
	not whispery	
3. feminine	fast	.21
	not whispery	

Table 5. Multiple regression analyses levels IV and V.

4.3.2. Speaker gender and voice quality (I and IV). The hypothesis is that the sex of the speakers and their self-assessed gender identity are expressed in their voice quality. This relation was analysed by means of two stepwise linear multiple regression analyses with the two voice quality factors (prosody and phonation) as criterion variables and sex of the speaker and their gender identity as predictor variables (Table 6). The analysis yields a model for the factor 'phonation'. However, inspection of the graphical presentation of the model showed that the significance of the relation was caused by a few outliers and did not present a regular pattern. The scores for the factor 'prosody' could not be predicted from the gender characteristics.

Factor	Model	R^2
1. prosody	—	—
2. phonation	masculine traits	.09

Table 6. Multiple regression analyses levels I and IV.

The results of the multiple regression analyses were significant but not satisfactory. On an individual level two correlations between voice quality features and gender characteristics also reached significance: *precise* and masculine traits (-.43, p < .01) and *nasal* and feminine behavior (.31, p < .05). Self-assessed masculine traits are related to precise articulation, and self-assessed feminine behavior is related to nasal voice. The listeners did not differentiate between the voices of the male and female speakers (note that voice height was scored sex-specifically).

4.3.3. Speaker gender and gender attributions (I and V). The hypothesis is that the self-assessed gender characteristics of the speakers are communicated to the listeners through the voice. Stepwise multiple regression analyses with the gender attribution factors as criterion variables and the gender characteristics of the speakers as predictor variables did not produce any model at all. The gender attributions therefore cannot be predicted from the speakers' gender characteristics.

However, on an individual level several significant correlations did emerge (Table 7). Male speakers are judged significantly more technical, more masculine, less emotional and less feminine than female speakers. In addition, speakers whose self-assessment of feminine behavior is relatively high are judged as less self-assured. A scatter plot of this relation showed this was true predominantly for male speakers. All these correlations are in the expected direction.

The last two correlations in Table 7 are caused by the effect that female speakers scored significantly higher than male speakers on their self-assessment of feminine behavior. Furthermore, as described before, the feminine and masculine judgments also correlated significantly with the sex of the speaker.

the sex of the speaker x	emotional	*	-.37
the sex of the speaker x	feminine	**	-.92
the sex of the speaker x	technical	**	.38
the sex of the speaker x	masculine	**	.92
feminine behavior x	self-assured	*	-.33
feminine behavior x	feminine	**	.39
feminine behavior x	masculine	**	-.38

Table 7. Significant correlations between levels I and V.
** = p < .01, * = p < .05

5. CONCLUSION

The relatively high reliability scores show that phonetically naive listeners are able to consistently describe voice quality features of voices and to consistently attribute gender-related personality traits to speakers on the basis of their voices. Furthermore, a multiple regression analysis revealed that the gender attributions are related to the voice quality descriptions. The most important result was the relation between 'dynamic' gender-related attributions (*enterprising*, *self-assured*, *adventurous*, and *independent*) and several voice features (*fast*, *relaxed*, *precise*, *rough*, *varied*, and *not whispery*). The results are comparable to those of Van Bezooijen [10], in which a lively manner of speaking (e.g. pitch variability) was strongly associated with among other things *activity* and *self-assured*.

The results further show that the correlations between the sex of the speaker and their self-assessed gender identity on the one hand, and the voice quality descriptions or gender attributions on the other hand, were not very large.

On the basis of these results and earlier published results on the relations between levels I, III, and V [1], and levels I and II [2], we expect that in the research model as a whole, the production side will predominantly yield relatively weak relations between speaker gender on the one hand and voice quality descriptions and gender attributions on the other hand. In contrast, on the perception side of the model relatively strong relations are expected between the voice quality descriptions and the gender attributions.

NOTES
1. We thank Jos Pacilly, Leiden University, for developing this program.

REFERENCES
[1] Biemans, M. 1998. Production and perception of gendered voice quality. To appear in the proceedings of *Perceiving and performing gender*. Kiel, Germany.

[2] Biemans, M. 1998. The effect of biological gender (sex) and social gender (gender identity) on three pitch measures. In Van Bezooijen, R. and Kager, R. (eds.), *Linguistics in the Netherlands 1998*, 41-52. Amsterdam: John Benjamins publishing company.

[3] Brown, B.L., Strong, W.J. and Rencher, A.C. 1974. Fifty-four voices from two: the effects of simultaneous manipulations of rate, mean fundamental frequency, and variance of fundamental frequency on ratings of personality from speech. *Journal of the Acoustical Society of America*, 55, 313-318.

[4] Graddol, D. and Swann, J. 1989. *Gender voices*. Oxford: Blackwell.

[5] Henton, C. 1989. Fact and fiction in the description of female and male pitch. *Language and Communication*, 9, 299-311.

[6] Henton, C.G. and Bladon, R.A.W. 1985. Breathiness in normal female speech: inefficiency versus desirability. *Language and Communication*, 5, 221-227.

[7] Pittam, J. 1987. Listeners' evaluations of voice quality in Australian English speakers. *Language and Speech*, 30, 99-113.

[8] Scherer, K. 1978. Personality inference from voice quality: the loud voice of extroversion. *European Journal of Social Psychology*, 8, 467-487.

[9] Tielen, M.J.T. 1992. *Male and female speech. An experimental study of sex-related voice and pronunciation characteristics*. PhD-thesis University of Amsterdam, the Netherlands.

[10] Van Bezooijen, R. 1988. The relative importance of pronunciation, prosody, and voice quality for the attribution of social status and personality characteristics. In Van Hout, R. and Knops, U. (eds.), *Language attitudes in the Dutch language area*, 85-103. Dordrecht: Foris Publications.

[11] Van Bezooijen, R. 1995. Sociocultural aspects of pitch differences between Japanese and Dutch women. *Language and Speech*, 38, 253-265.

[12] Van Rie, J. and Van Bezooijen, R. 1995. Perceptual characteristics of voice quality in Dutch males and females from 9 to 85 years. *Proceedings of the XIIIth International Congress of Phonetic Sciences*, 2, 290-293.

[13] Willemsen, T.M. and Fischer, A. 1994. De Nederlandse sekse-identiteit vragenlijst (The Dutch gender identity questionnaire). *Nederlands Tijdschrift voor de Psychologie (Netherlands Journal of Psychology)*, 52, 126-130.