EVALUATION OF VOICE AND PRONUNCIATION CHARACTERISTICS OF MEN AND WOMEN

Mirjam T.J. Tielen and Florien J. Koopmans-van Beinum

Institute of Phonetic Sciences, University of Amsterdam,
Herengracht 338, 1016 CG Amsterdam, The Netherlands.

ABSTRACT

In this paper results are presented of an extensive listening experiment on the evaluation of male and female voice and pronunciation characteristics. Thirty male and thirty female speakers from three profession categories were recorded while reading aloud several texts. Three main research questions were involved: 1. Are voice and/or pronunciation of men and women evaluated differently? 2. Do listeners' judgments and subjects' opinions about these characteristics reveal similar results? 3. Can speakers from different professions be distinguished by voice and pronunciation cues only?

1 INTRODUCTION

Clear differences in acoustic characteristics exist between male and female voices. An interesting question in this connection is which perceptual characteristics would be related more to male voices and which to female voices. A study by Kramer [2] revealed that some perceptual characteristics were more associated with female speech (e.g. gentle, melodious), while other characteristics were associated more with male speech (authoritative, loud). In the present experiment, it was tested to what extent male and female voice and pronunciation would be evaluated differently. Judgments based on actual presentation of voices and subjects' opinions were compared. All evaluation scores were collected by means of semantic scales [3]. Another question was whether voices of different professions would be evaluated differently. If so, this would imply that listeners are able to distinguish voices with respect to profession (see also [4]).

2 METHODS

2.1 Speakers and listeners

Thirty male and thirty female representatives from three profession categories (nurses, managers and information agents) were selected. These particular speaker groups have been chosen, because these groups differ clearly from one another with respect to the number of men and women working in these professions, and because speech is an important aspect of the work in all three categories. Twenty male and twenty female students of the University of Amsterdam (all native speakers of Dutch) participated as listeners in the experiment.

2.2 Stimuli and Recordings

The speakers were asked to read aloud text passages taken from actual speech. Three texts dealt with topics associated with the three professions. An additional text was included which dealt with a neutral topic with respect to the speaker groups. The sixty speakers were recorded at various places. They were allowed to prepare the texts in the way they desired; also, they were free in choosing their own tempo and intonation. At the end of the recordings, the speakers gave their opinion about their own voice and pronunciation by means of the already mentioned rating instrument (results in 3.3).

2.3 Perceptual evaluation

Voice and pronunciation were evaluated by means of the semantic 'twin scales' (seven pairs of related notions) as developed for Dutch by Fagel et al. [1].

In addition, four new scales were included that were supposed to differentiate between the profession categories, if the professions would differentiate at all. The English translations of the Dutch scales as used in the present experiment are shown in Table 1. The four scales at the bottom of the table are the additional ones.

2.4 Experimental procedure

The listeners were instructed to evaluate the voices by means of the eighteen scales (results in 3.1). The seven texts as well as the sixty speakers were randomized. The listening sessions were performed at the Language Centre of the University of Amsterdam, where listening facilities were available that allowed for selective presentation to listeners individually and simultaneously. During the listening sessions, the forty listeners were also asked to identify the profession of the speakers presented by choosing from a list of six possibilities (results in 3.4).

Apart from evaluating the voices, the listeners were also asked to give their opinion about typical voice and pronunciation characteristics of men and women in six profession categories, including the three already mentioned (results in 3.2). This task was also performed by means of the same rating instrument.

3 RESULTS

3.1 Evaluation of speakers' voice and pronunciation

Mean scale scores were derived for all three speaker groups for men and women separately (see Table 1). It appeared that the differences between male and female voice and pronunciation are not very large, although except for scales no. 1,9,11,12 and 18, the scales appeared to be significantly different (t-test; sign. level was set to 0.003, because of the repetition of t-tests). Not surprisingly, the largest differences are found for the scales 'high-low' and 'shrill-deep'. Smaller differences are found for the scale 'dull-clear', with female voices considered to sound clearer than male voices. Furthermore, female nurses and managers were evaluated as more monotonous than male nurses and managers. Some scales differentiated the profession categories. Managers were evaluated as speaking in a little more polished and cultured way. Factor analyses were performed on the correlations between the scales in order to look for the underlying patterns of relationships between the data. Analyses performed on the male and female data separately, revealed that four factors explained about 50% of the total variance in rather well interpretable dimensions that can be characterized as 'Appreciation quality of the voice', 'Personality evaluation', 'Pronunciation quality' and 'Pitch' respectively, with some minor differences between the two sexes. The scales 'broad-cultured' and 'pleasant-unpleasant' attained the highest communality estimates, which implies that these scales are responsible for a considerable part of the variance.

3.2 Opinions about typical voice and pronunciation characteristics

The mean scale scores obtained by the non-auditively based judgments reveal that there a smaller number of assumed differences between the voices of the two sexes exist in comparison to the the voices of the three profession categories. Only the scales 'high-low' and 'shrill-deep' differed significantly for male and female voices.

Most extreme scores were found for the scales 'slovenly-polished' and the related scale 'broad-cultured'. Managers (male as well as female) were supposed to possess the highest degree of culture and polishment in their pronunciation. Also, rather extreme mean scores are found for attributes like powerful, authoritative and business-like with respect to managers. One more appealing finding was the high mean score for male information agents with respect to melodiousness.

3.3 Evaluation of voice and pronunciation by the speakers themselves

The mean scores as given by each of the speakers with regard to their own voice and pronunciation also tended to the centre of the interval. Nevertheless, female information agents judged their pronunciation as more polished and cultured than the other speaker groups.

Table 1.
Mean scale scores for female (F) and male (M) nurses (20), managers (20), and information agents (20) as evaluated by all (40) listeners over all text presentations. In the last two columns the overall means for female and male speakers are presented. The 18 scales used were seven-point interval scales.

SCALES		NURSES		MANAGERS		INF.AGENTS		ALL	
		F	M	F	M	F	M	F	M
1. slovenly	polished	4.58	4.39	4.87	4.75	4.35	4.63	4.60	4.59
2. broad	cultured	4.28	4.42	4.67	4.59	4.08	4.65	4.34	4.55
3. high	low	3.76	4.98	3.90	5.05	3.50	4.65	3.72	4.89
4. shrill	deep	3.88	4.80	3.99	4.90	3.47	4.49	3.78	4.73
5. dragging	brisk	4.34	3.89	4.36	3.97	4.19	4.65	4.30	4.17
6. slow	quick	4.14	3.56	4.10	3.70	4.06	4.33	4.10	3.86
7. husky	not husky	4.13	4.37	4.24	4.28	4.31	4.47	4.23	4.37
8. dull	clear	4.59	4.05	4.78	4.06	4.75	4.30	4.71	4.14
9. weak	powerful	4.25	4.21	4.56	4.38	4.32	4.48	4.38	4.36
10. soft	loud	4.14	3.97	4.31	3.94	4.51	4.18	4.32	4.03
11. ugly	beautiful	4.23	4.02	4.35	4.19	3.72	4.15	4.10	4.12
12. unpleasant	pleasant	4.49	4.29	4.63	4.44	3.96	4.43	4.36	4.39
13. monotonous	melodious	4.40	3.95	4.66	4.09	4.34	4.25	4.47	4.10
14. expressionless	expressive	4.37	3.91	4.68	4.08	4.37	4.22	4.47	4.07
15. severe	sweet	4.45	4.14	3.96	3.93	4.07	3.96	4.16	4.01
16. business-like	emotional	4.07	3.74	3.59	3.53	3.99	3.57	3.88	3.61
17. vulgar	distinguished	3.87	4.06	4.31	4.40	3.80	4.16	3.99	4.21
18. timid	authoritative	3.86	3.97	4.40	4.17	4.14	4.23	4.13	4.12

Female managers and information agents scored their voices as more expressive than the other speaker groups. Male speakers regarded their own voices less husky than the female speakers did.

3.4 Identification of profession by voice cues alone

Most listeners reported that identification of the profession of a speaker was a difficult task. The alternative options of 'shop assistent' and 'teacher' (categories that were actually not present in the speaker groups) appeared to be options that were rather frequently chosen. It appeared that female nurses were identified correctly more often than male nurses. Moreover, female nurses were scarcely confused with female managers in contrast to male nurses with male managers. Male information agents were classified most often as teachers, whereas female information agents received most scores on the categories of nurse and shop assistent. The significance of the different scores was tested in an analysis of variance.
The Anova analysis (mixed model with repeated measures on the factors

Table 2.
Results of the analysis of variance (mixed model; repeated measures) on the correct profession identification scores. Sl=Sex of listener; Ssp=Sex of speaker; Psp=Profession of speaker.

FACTOR	DF	SS	F RATIO	SIGN. OF F
Sl	1, 38	10.37	1.48	.23
Ssp	1, 38	27.26	5.94	.02
Psp	2, 76	143.52	11.49	.00
Ssp * Psp	2, 76	136.41	13.87	.00
Sl * Ssp * Psp	2, 76	4.25	.43	.65

'speaker sex' and 'profession of speaker') gave rise to the results as presented in Table 2. From that table it can be seen that sex of speaker was one of the differentiating factors.
Also, speakers with different professions were identified differently and the interaction between sex of speaker and profession of speaker also reached the level of significance, which implies that male and female nurses, managers and information agents received different scores. From the differences in percentages between the three speaker groups, it appeared that the manager scores are higher when manager voices had been presented. The same holds for the nurse scores. Although not tested in the Anova, the influence of text condition on the profession appointments appeared to be very clear as well. The scores on each of the categories were increased when the text content fitted with that particular profession.

4 DISCUSSION

Although considerable differences existed between the individual speakers of each group as evaluated by the listeners, in general the differences between male and female speakers as well as between managers, nurses and information agents appeared to be rather small. More differences were appointed with respect to male and female speakers if voice and pronunciation characteristics were evaluated without actual presentation of voices. These supposed characteristics were also more extreme. The finding that the differences are considered to be larger than were actually found, may point to the existence of prejudices. The differences between 'high-low' and 'slow-quick' were enlarged for male and female managers. Differences in 'high-low', 'ugly-beautiful' and 'monotonous-melodious' were enlarged for male and female information agents. The differences between the three professions were even more enlarged; managers are supposed to speak e.g. very polished and with authority in comparison with the other groups. Evaluation by the speakers themselves revealed not such clear group differences.
Identification of profession on the basis of someone's voice was far from perfect, but neither a random choice.

5 REFERENCES

[1] Fagel, W.P.F., Herpt, L.W.A. van, and Boves, L. (1983), "Analysis of the perceptual qualities of Dutch speakers' voice and pronunciation", *Speech Communication 2*, 315-326.
[2] Kramer, Ch. (1977), "Perceptions of female and male speech", *Language and Speech 20*, 151-161.
[3] Osgood, C.E. and Suci, G.J. (1955), "Factor analysis of meaning", *J. of Exp. Psychology 50*, 325-338.
[4] Tielen, M.T.J. (1990), "Perception of the voices of men and women in relation to their profession", *Proc. ESCA Workshop on speaker characterization*, 192-197.