SOCIOPHONETIC ASPECTS OF DUTCH CLEFT-PALATE SPEECH

ALDA VAN ERP

Dept. of Applied Mathematics and Signal Processing,
Dr. Neher Laboratories Netherlands PTT,
Leidschendam, the Netherlands

ABSTRACT

The results of a rating experiment are presented and related to 18 phonetic variables obtained from a panel of phonetic experts. Recorded, non-emotional speech fragments spoken by ten adult male Dutch cleft-palate speakers were investigated. The verbal content was controlled in all speech fragments. In the rating experiment the cleft-palate speech fragments were rated by two groups of listeners (of whom 30 had had a training in speech therapy and 30 had not) on 19 speech scales and 15 scales pertaining to social status, social attractiveness and competence.
The results show which variables cause listener group effects. Additionally, the relation between speech ratings and personality/social ratings are displayed and compared for the two listeners' groups. Moreover, the listeners' ratings are related to the (expert) phonetic variables.

INTRODUCTION

Cleft-palate speech is the speech produced by someone who has (had) a cleft palate. It is pathological in the sense that it sounds obviously deviant from speech that falls within the range that is accepted as 'normal' by the speakers in a particular speech community. It deviates from 'normal' on a number of vocal aspects pertaining to articulation, phonation, resonance and prosody.
Vocal aspects may be used by listeners to infer information about characteristics of the speaker. For example, when you talk to someone whom you have just met for the first time, it may happen that you get a first impression of the other which is based on how the other speaks rather than what the other says. Although it may happen that such inferences are not in keeping with the reality, they play an important role in (first) impression formation.
It is not clear to what extent typical cleft-palate vocal characteristics contribute negatively to the impression people get of the speaker in question. The present paper attempts to answer this question. To this end perceptual descriptions of the relevant vocal characteristics were related to inferences about speaker characteristics that are based only on vocal information.
The process of attributing speaker characteristics from only vocal aspects appears to include not only conclusions about psychological and social aspects, but also about physical aspects of the speaker's identity such as his or her sex, age, height, weight, physique and state of health [2]. The study reported on here does not deal with inferences about physical aspects of the speakers. It deals with personality and social aspects on the basis of vocal aspects. Moreover, the personality and social judgments in this study are obtained from listeners only; they are not obtained by means of tests.

METHOD

At the base of the research reported on here are running speech fragments from a sample of Dutch cleft-palate speakers. Firstly, the vocal aspects of these fragments have been phonetically described. This was done analytically by a panel of experts. Secondly, the same vocal aspects were judged in an associative fashion. This was done in a rating experiment by various relatively large groups of listeners. Thirdly, the speech fragments were used for obtaining associative, inferential judgments about the speakers' personality and social aspects. This was done in a rating experiment as well, by the same listeners that also rated the vocal aspects in an associative fashion. Thus, the cleft-palate speech samples were described on three levels. Analogous to the lens model [1], these three levels of description can be referred to as 'distal', 'proximal', and 'attributional'. On the distal level is the phonetic description of the vocal aspects; on the proximal level the associative description of the vocal aspects; and on the attributional level the associative description of the personality and social aspects. The rating experiment was designed in such a way that the proximal and the attributional levels were distinguished.

Speech material

The data base consisted of recorded prose passage renderings by ten different cleft-palate speakers. The prose passage was an emotionally neutral reading text, and yielded more than one minute of running speech per speaker, the verbal content being controlled. The speakers were male, had slight South-Eastern Dutch accents, and were aged between 17 and 48 years.

The phonetic description

A combined approach was followed: Both a segmental and a nonsegmental description were made of the speech material (see above) by four experts. In addition, 12 linguistically trained judges were used for obtaining intelligibility scores. These were based on nonsense sentences read aloud by the ten cleft-palate speakers.

The segmental description indicated which phonemes were pronounced deviantly and how often each of the following typical errors were made: (1) fronting of the place of articulation,(2) backing of the place of articulation,(3) glottal stop, (4) nasal emission,(5) nasal explosion, and (6) denasality.

The nonsegmental description consisted of ratings on 33 vocal parameter scales. By means of scalar degrees it could be indicated for each parameter whether the deviation from a predefined neutral point was either 0, 1, 2, or 3. The ratings are a mix of quality and quantity. This means that if a particular vocal effect is very strong when it is present it would be rated as 3 if it occurred relatively often. However, if it occurred relatively rarely it would receive a lower score. The scales pertain to: (1) supralaryngeal features (concerning the lips, jaw, tongue tip, tongue body, the velopharyngeal mechanism, the pharynx, as well as supralaryngeal tension and precision of articulation), (2) laryngeal features (concerning phonation type and laryngeal tension), and (3) prosodic features (concerning pitch, loudness, and temporal structure). The intelligibility scores were expressed in percentages of syllables that were reported correctly, averaged over the 12 judges.

The associative description

For the associative description, the speech material that was phonetically described was presented to 60 female listeners. Their mean age was 22 years, ranging from 20 to 26 years. 30 listeners were students from the college of speech therapy training in Nijmegen ('TRAINED'). The other 30 were students enrolled in the Faculty of Arts of the University of Nijmegen, but not in language courses ('UNTRAINED'). The listeners were born and raised in the South-Eastern part of the Netherlands and were therefore accustomed to a South-Eastern Dutch accent. The rating experiment took place in a language laboratory with individual booths. The listeners were presented with the recorded speech material via headphones. They did not know any of the speakers nor did they know what the speakers looked like. The listeners were asked to rate each bi-polar (7-point)scale on the rating sheets they got in front of them. The meanings of the scale positions was explained to them in the written instructions. The instructions also encouraged them to give their first impressions. In fact, they were only given approximately 3 seconds per scale to respond. The rating scales were divided into two categories. One category contained scales pertaining to vocal aspects ('speech scales'), the other contained

scales'), the other contained scales pertaining to personality and social aspects. The two types of scales were rated in separate sessions.

Twelve speech scales refer to more or less general vocal aspects that pertain to articulation, phonation type and prosody **(viz. standard, precise, intelligble, good reading performance; bright, creaky; high-pitched, varied, expressive, loud, quick, smooth).** Seven speech scales refer to pathological vocal aspects **(viz. nasal, with a blocked-up nose, snorting, snoring, glottalised, hoarse, lisping).** In addition, there was one question with a dichotomous response category, namely: "Do you think speech therapy is required? yes/no".

The personality/social scales refer not so much to the classic Evaluation, Potency and Activity dimensions, but rather to dimensions that were considered to be more relevant in connection with the social acceptability of people with a speech defect. Therefore, they refer to social status, (social) competence and social attractiveness **(viz. of high social status, highly ambitous, with qualities of leadership, self-confident, reliable, intelligent, suited for public speaking, strong-willed, careful, interested, friendly, warm-hearted, spontaneous, cheerful, modest).**

RESULTS AND DISCUSSION

Firstly, I will discuss the outcomes of t tests based on the associative ratings. This is done, separately for the speech ratings and the personality/social ratings, to determine whether there is any difference between the ratings of the two groups of listeners (viz. TRAINED versus UNTRAINED). Subsequently, the relations between associative speech ratings and personality/social ratings will be discussed and compared for the two groups of listenerss. Finally, the relations between the analytic speech description (esp. of the pathological aspects) will be related to the associative ratings. Again, the groups of listeners will be compared.

Before the associative ratings were subjected to t tests, interrater reliabilities (Cronbach's alpha) were computed, separately for each scale, and separately for the two groups of listeners. For the speech scales, these coefficients generally appeared to be satisfactorily high (>.80) and comparable for the two groups of listeners. The only conspicuous differences between the two listener groups occurred for a few pathological speech scales. Firstly for **not sniffing- sniffing, not snoring-snoring, not hoarse-hoarse, and not creaky-creaky,** where the reliabilities of the UNTRAINED listeners were comparatively low, ranging from .44 for **not sniffing-sniffing** to .74 for **not creaky-creaky** while the reliabilities of the TRAINED listeners were higher than .90. These differences were due to the fact that the TRAINED listeners indicated differences between the various speech fragments more clearly than the UNTRAINED listeners. And secondly there was a conspicuous difference for **not with a blocked-up nose-with a blocked-up nose,** where

the reliability for the TRAINED listeners was lower than for the UNTRAINED listeners was lower than for the UNTRAINED listeners (viz. 71 versus .91). This was mainly due to the fact that the TRAINED listeners did not seem to agree much among themeselves with respect to the scale positions they assigned to individual speech fragments. As for the personality/social scales, there appeared to be no differences between the two listeners groups that are worth mentioning. Twelve scales were rated very reliably > .91). Three scales (viz. unfriendly-friendly, conceited-modest, and unreliable-reliable) were rated less reliably, with coefficients ranging from .69 to .80. In addition, it appeared that in general the personality/social ratings were less extreme than the speech ratings. This could mean that the listeners are rather careful in attributing personality/social characteristics to speakers on the basis of only their speech.

In subsequent analyses use was made of the mean of the ratings of 30 listeners on each of the 19 speech scales and 15 personality/social scales respectively, for each of the ten cleft-palate speech fragments.

The t test results revealed that statistically the ratings of the two groups of listeners (averaged over 30 listeners and 10 speakers) were equally high. Therefore, the conclusion is that there is no effect for groups of listeners, neither for the speech ratings nor for the personality/social ratings. With respect to the speech ratings this could mean that also in a normal social context the cleft-palate speech aspects are just as salient for laymen as for speech therapists. With respect to the personality/social ratings this means that apparently listeners are rather consistent in the attribution of personality/social characteristics on the basis of someone's speech only.

In order to determine whether there are any relations betwecn associative ratings of vocal aspects and associative ratings of personality/social aspects, product-moment correlations were computed. Firstly this was done, separately for the two groups of listeners, for general ratings (i.e. ratings that are not only averaged over listeners but also over scales). This general correlation was .80 for the TRAINED listeners and .88 for the UNTRAINED listeners.

In order to be able to investigate the relation between associative speech ratings and associative personality/social ratings in more detail, correlations were subsequently computed between every speech scale and every personality/social scale, separately for the two groups of listeners. From the results it appeared that, for both groups of listeners, the more general (i.e. not-pathological) speech ratings correlate significantly with personality/social ratings far more often than do pathological speech ratings. Speakers that were judged to speak little **varied, expressive, precise, smooth, clear, loud,** and who were judged to have a **bad reading performance** were judged less positively on most personality/- social scales than were speaker

that were judged to speak **varied, expressive, precise** etcetera. It should be noted, however, that this finding may be an artefact of the speech material that, was used (viz. a reading text). As for the pathological speech ratings, there were some remarkable differences between the two groups of listeners. For the UNTRAINED listeners there were significant correlations between ratings of a number of pathological speech aspects (viz. **nasal, snorting, snoring)** and judgments about certain personality/social aspects (viz. **unsuited for public speaking, without qualities of leadership, modest, wavering).** For the TRAINED listeners the judged presence of pathological speech aspects did not correlate significantly with personality/social judgments. This difference between the two groups of listeners is important insofar that it seems to indicate that a layman is more inclined to attribute certain less positive personality/social characteristics to speakers with obvious cleft-palate speech defects than a speech therapist.

Before the analytic description of the vocal aspects were related to the associative ratings, statistical analyses were carried out – for the nonsegmental description and the intelligibility scores – in order to make sure that the ratings were reliable, and that the various parameters were relevant to the aim of the study. The reliability of the means of the scores was assessed by means of Cronbach's alpha.

It appeared that for the nonsegmental parameters this coefficient ranged from .06 for protrusion of the lower jaw to .88 for speech rate. Only values higher than .75 were considered to be satisfactorily high. Consequently, only ten out of the 33 nonsegmental parameters were considered to have been reliably rated, namely: nasality, nasal emission, precision of articulation, whisperiness, creakiness, pitch mean, pitch range, loudness mean, interruptedness, and rate. The parameters that were rated most severely, averaged over the ten speakers, were nasality, nasal emission and whisperiness. In order to assess whether the ratings on these scales varied as a function of the speakers, the ratings on each of these scales were subjected to separate analyses of variance with two fixed factors namely 'speakers' and 'raters' (level of significance= 5%). It appeared that the factor 'speakers' was significant for all ten scales. Inspection of the mean ratings (N=4) for each these parameters revealed that this was not caused by just one or two speakers who had received extreme ratings while the other speakers were rated neutral.

For the average intelligibility scores (N=12) the reliability coefficient, averaged over ten speakers, was .85. The scores for the individual speakers ranged from 80% to 98%. However, for nine out of ten speakers the range was between 92% and 98%. This points to a ceiling effect. Additionally, in order to examine how the 18 different analytic variables (i.e. 7 segmental, 10 nonsegmental, and 1 intelligibility variable) were related, product-moment correlations were

computed. There were nine significant correlations and in only five cases the correlation was so high that more than half of the variance in one variable was accounted for by another variable. This concerns the following variables: (nonsegmental) nasality and (nonsegmental) nasal emission (r= .71), (segmental) nasal emission and (nonsegmental) nasal emission (r= .79), (segmental) nasal explosion and (nonsegmental) loudness mean (r= .71), (segmental) glottal stops and intelligibility (r= -.87), and precision of articulation and pitch range (r= .71). Because neither of these correlations are extremely high, it was decided to relate all 18 phonetic variables to the associative ratings.

To determine the relations between the phonetic variables and the associative ratings, product-moment correlations were computed, separately for the two groups of listeners.

In the first place, this was done between the phonetic and the associative descriptions of the vocal aspects. Correlations between the following parameters were significant, for either one or both of the groups of listers. Correlations higher than |.63| are significant. The height of the correlations indicates the validity of the associative ratings of the speech aspects. The first correlation is of the UNTRAINED listeners; the second of the TRAINED listeners.

1. (Segmental) fronting of place of articulation with lisping (.84, .81)
2. (Segmental) nasal explosion with snoring (.48, .66)
3. (Nonsegmental) nasality with nasal (.92, .83)
4. (Segmental) nasal emission with snorting (.40, .81)
5. (Nonsegmental) nasal emission with snorting (.64, .90)
6. Intelligibility with intelligible (.54, .69)
7. (Nonsegmental) whisperiness with hoarse (.88, .90)
8. (Nonsegmental) precise articulation with precise (.86, .83)
9. (Nonsegmental) rate with quick (.82, .80)

Apparently, for snoring, snorting, and intelligible the associative ratings of the TRAINED listeners are clearly more valid than those of the UNTRAINED listeners. For lisping, nasal, hoarse, and precise there is practically no difference.

In the second place, this was done between the phonetic description of the vocal aspects and the associative description of personality and social aspects. Correlations between the following parameters were significant for either one or both of the listener groups. Their height indicates the strength of the relationship between 'true' vocal characteristics (i.e. vocal characteristics as described analytically by experts who were trained to do the job) and inferred personality/social characteristics. Again the first correlation is of the UNTRAINED listeners; the second is of the TRAINED listeners.

1. Nasality with suited for public speaking (-.61, -.66)
2. Whisperiness with intelligent (-.71, -.68)
3. Whisperiness with of high social status (-.66, -.69)
4. Whisperiness with with qualities of leadership (-.62, -.66)
5. Whisperiness with suited for public speaking (-.62, -.64)
6. Rate with spontaneous (.56, .64)
7. Rate with modest (-.71, -.76)

Apparently, these correlations are much the same for the two listener groups. In addition, it appeared from the results that precise articulation and pitch range correlate significantly with every personality and social scale except for suited for public speaking, self-confident, and modest. Their correlations ranged from .68 (e.g. for intelligent) to .91 (for careful) and were comparable for both groups of listeners.

CONCLUSION

From the results it appeared that there are obvious relations between some pathological vocal characteristics (viz. nasality and whisperiness) and negative ratings on some personality/social characteristics pertaining to social competence. These relations were more or less equally strong for TRAINED and UNTRAINED listeners. However, it also appeared that for more general (i.e. not-pathological) vocal characteristics (viz. rate, precision of articulation, and pitch range) the relations with rating on personality and social characteristics are more pervasive. Moreover, the results strongly suggest that these five vocal characteristics mediated in the attribution process. Admittedly, whether these vocal characteristics are actually used for attributing personality and social characteristics of the speakers in a normal social context would have to be investigated in a more realistic setting.

REFERENCES

[1] Brunswik, E. Perception and the representative design of psychological experiments. Berkeley and Los Angeles, California: University of California Press, 1956.
[2] Laver, J. & Trudgill, P. Phonetic and linguistic markers in speech. In: Scherer, K. & Giles, H. (eds.), Social markers in speech. Cambridge: Cambridge University Press, 1979.