PHONETIC CRITERIA OF ATTRACTIVE MALE VOICES

Vivien Zuta

Phonetisches Institut, Johann Wolfgang Goethe Universität, Frankfurt am Main
Zuta@fias.uni-frankfurt.de

ABSTRACT

This paper investigates the attractiveness of German male voices for female listeners. The primary finding of this study is that deep voices are not necessarily attractive for women. Even voices with a middle or high fundamental frequency can be judged as attractive.

Furthermore, the analysis shows that there is a combination of various parameters, which are responsible for leaving positive or negative impressions to the listener.

Keywords: male, voice, attractiveness, speaker, listener, German

1. INTRODUCTION

An investigation on the extent to which physiological characteristics of a speaker can be judged from his voice was conducted by van Dommelen and Moxness [3]. They found significant correlations between estimated height/weight and actual height/weight only for male speakers. An important factor for these correct estimations was the sex of both speaker and listener. Dominantly male listeners were able to estimate actual heights/weights of the speakers. Neither males nor females were able to estimate the height/weight of female speakers.

Another investigation was performed by Bruckert et al. [4] who found that females were consistent in their judgment of pleasantness and in their height, weight and age estimations. Bruckert et al. concluded that pleasantness judgments were based mainly on intonation. Women tend to prefer voices with raising pitch. Another result was that female listeners were able to correctly estimate age and weight but not height. For the age estimation women would use lower-frequency formants with little formant dispersion as a sign for an older age. Bruckert et al. however were not able to determine which acoustic parameters were used to estimate the subjects weight correctly.

In this paper we investigate several acoustic parameters for their correlation with the attractiveness females assign to male voices.

2. MATERIAL

The recordings for the experiment were conducted using a set of six German male speakers who were requested to narrate the fairy tale "Little Red Riding Hood". (Little Red Riding Hood was chosen because of its widespread familiarity.) The task for the female listeners was to judge in reference to the sound they were listening to but not in reference to the content of what the male speakers were saying. It was ensured that everybody who took part in the investigation was familiar with the fairy tale. Speakers were told to narrate the story in a natural way and they were recorded in a sound treated environment with professional equipment. The recordings were played via loudspeakers to females, who were requested to fill out a questionnaire in which they were asked to:

- describe the assumed appearance of the speaker
- judge the level of education
- state if they thought that the speaker was attractive based on his voice.
- judge the quality of voice (by given adjectives)
- tell if they would say the voice itself sounds attractive
- name a male person, which, in their opinion, has an attractive voice and state the relationship that this male person has to themselves.

The speakers were divided into two categories, chosen by the evaluated data. They were labelled 'more attractive sounding voices' and 'less attractive sounding voices'. The measured data were compared within the groups and between the two groups. Additionally cardinal vowels embedded in common German words were recorded.

3. METHOD AND RESULTS

The analyzed and measured parameters were:

- Fundamental frequency (F0)
- Standard deviation of F0
- Coefficient of variation of F0
- Pauses
- Hesitations
- Long term average spectra (LTAS)
- Speech rate
- Speed of articulation
- First three formant frequencies of [a:], [e:], [i:], [o:], [u:]

The acoustic analysis was conducted manually and with help of the speech processing software PRAAT. For the running speech analysis the whole recordings of the retold fairy tale "Little Red Riding Hood" was examined for each speaker. The formant frequencies were measured at the center of the cardinal vowels of separate words.

3.1. Analysis of the fundamental frequency (F0), standard deviation of F0 and the coefficient of variation of F0.

The assumption 'attractive male voices must be deep' upon which this investigation started was not confirmed. Within the group of the more attractive sounding voices the values of the fundamental frequency were reaching up to 134Hz. This was analogue to the result in the group of less attractive sounding voices. The deepest voice of this group had a fundamental frequency of 101Hz, which can be seen as a normal or even deep voice. This result leads to the conclusion that deep voices do not necessarily leave a positive impression on the (female) listener (as opposed to the assumption we started working on).

The result of the analysis of the fundamental frequency showed that a classification of the attractiveness of a voice cannot be made only by the F0 value alone, which also gets clear when we look at the average F0 values of both groups. The more attractive voices have an average F0 of 112Hz and the average F0 for the less attractive voices is 110Hz. It will be shown that there are additional important parameters.

The examination of the standard deviation of F0 showed that modulation of F0 increases attractiveness whilst monotonous voices are disfavoured. The group of less attractive sounding voices has low standard deviations between 16Hz and 21Hz whereas the value of this parameter was considerably higher for the group of the more attractive sounding voices (between 19Hz and 29Hz).

The analysis of the coefficient of Variation of F0 (standard deviation divided by the mean) showed similar results: a small value for the coefficient goes along with a higher chance of a positive estimation of the speakers voice.

	mean F0 [Hz]	st. dev. [Hz]	
Speaker 1	117	21	
Speaker 2	101	17	LAV
Speaker 3	113	16	
Speaker 4	134	29	
Speaker 5	108	26	MAV
Speaker 6	94	19	

(Increasing attractiveness ↓)

Table 1: Fundamental frequency and standard deviation of LAV (less attractive voices) and MAV (more attractive voices).

	Average mean F0 [Hz]	Average St. dev. [Hz]
LAV	110	18
MAV	112	24

Table 2: Average fundamental frequency and average standard deviation of LAV (less attractive voices) and MAV (more attractive voices)

3.2. Analysis of the formant frequencies

The comparison of the first Formants (F1) of the speakers does not reveal a significant pattern. The second Formants (F2), however, showed a tendency: the average F2 of the first group (less attractive sounding voices) is 53Hz lower than the average F2 of the group with the more attractive sounding voices. A possible reason for this result could be that F2 manifests a bigger effort in the continuous movement of the articulators to produce running speech and pronunciation of the second group. It is conceivable that this is connected with a higher activity of the lips and the

tongue and leads obviously to a higher acceptance for the average German female listener. However, a more detailed examination of this factor is a viable subject for further studies.

3.3. Analysis of long-term average spectra (LTAS)

Comparison of the long-term spectra of the less attractive voices, reveals a significant dip at around 2.8kHz, which is non-existent in the spectra of the other group. To figure out a reason for this notch a comparison with the depiction from Nolan's [1] investigation was made. The interpretation of Nolan's results leads to the assumption that, in this study, the less attractive voices contain a higher grade of nasality. The characteristic is audible and was picked up by some of the female listeners. From the results it is obvious that a voice quality, influenced by nasality, degrades the attractiveness of a voice.

A noticeable attribute for this analysis is that all of the less attractive sounding voices show this certain irregularity in their spectra, whereas none of the attractive sounding voices show characteristics such as the mentioned dip round 2.8kHz. Therefore can be assumed that the attraction of a voice is conditional to the naturalness of the sound. The fewer parameters of articulation are being changed or altered in relation to a "standard voice" (what Hollien [2] calls a "0-Voice"), the less the standard pronunciation and flow of articulatory movements is influenced the more pleasing a man's voice seems to be.

3.4. Analysis of pauses

In this experiment pauses were defined as silent periods in the speech flow. A pause was detected where the silent period was not interrupted by hesitations and/or speech. Pauses were measured and counted

3.4.1. Analysis of the average length of pauses

It is striking that most speakers have an average length of pauses of 0.6s. Two speakers, whose voices belong to the group of the less attractive voices, have differing values, which were 0.3s and 0.8s.

Again we find an agreement with the previous results, which showed that a voice leaves a more attractive impression when it is close to a standard voice, with as little as possible variance in the parameters of an undisturbed articulation.

3.4.2. Analysis of the pause-free interval

The measurements of the pause-free intervals gave additional insights. The average pause-free interval is about 0.7s longer in the group of less attractive voices than in the more attractive voices. Which means that the flow of words is longer than in the group of the more attractive voices. This indicates that a voice seems to be judged as more attractive when the distance between two pauses is not too long or too short.

This result counts of course only for the "normal" stream of speech. We cannot work on the assumption that a speech signal composed of almost only pauses would sound good or attractive for a listener.

	Pause-free interval in sec.	Average pause-free interval		
Speaker 1	3.0			
Speaker 2	3.7	3.7	LAV	
Speaker 3	4.3			
Speaker 4	2.9			
Speaker 5	3.7	2.9	MAV	
Speaker 6	2.2			

Increasing attractiveness

Table 3: Pause-free interval and average values of LAV (less attractive voices) and MAV (more attractive voices).

3.5. Analysis of hesitations

For this study the number and duration of hesitations were defined as 'filled-pauses' of the continuous speech flow. Furthermore the repetition of word beginnings was counted to the hesitations as well. The analysis of hesitations includes the examination of hesitations per minute, the average duration of hesitations during the whole recorded signal and the hesitation-free interval.

3.5.1. Results of the hesitations per minute

On average the quantity of hesitations per minute in the group of the less attractive voices is more than twice as high as the average number of hesitations per minute in the more attractive group (11 versus 5 per minute). It is noticeable that the lowest number of hesitations per minute was detected in the latter group (3 per minute) whereas the highest number was found in the first group of voices (12 per minute).

3.5.2. Analysis of the average length of hesitations

The length of hesitations is another influential factor. In analysing the results for the voice rated as most attractive and the voice stated as most unattractive we find again (see 3.5.1.) the lowest (attractive voice) and the highest (unattractive voice) values of 0.3 s and 0.8 s respectively.

The longer the duration of the hesitations the less attractive the voice is considered to be.

3.5.3. Analysis of the hesitation-free interval

The result of the study of this parameter shows that the average value of the hesitation-free interval in the group of those voices, which were characterised as less attractive, was only 4.9 s whereas the value for the attractive voices was 14.1 s.

A word flow, which is more often interrupted by hesitations, is judged as unattractive.

3.6. Analysis of speech rate

Analogue to Koremann [5] the results of the investigation of the speech rate show that it is not useful to count and calculate the syllables per minute without relating this attribute to different factors such as pauses and hesitations. In this study the conclusion was made by the examination of the syllables per minute. In direct comparison no significant difference was found. The average difference between both groups was on average only around 0.2 syllables per seconds (including pauses).

3.7. Analysis of speed of articulation

For the analysis of speed of articulation the same arguments as in 3.6. can be put forward even though some very small differences were found. The average difference between both groups lies around 0.4 syllables per s (excluding pauses). Closer examinations are definitely necessary in order to give a clear answer about the question if the speed of articulation is an important factor for the judgement of female listeners.

4. CONCLUSION

There are obvious and crucial parameters, which are important for the attraction of a male German voice.

- High-pitched voices can be attractive.
- Modulation of F0 increases attractiveness.
- Nasality degrades the attractiveness of a voice.
- Moderate length of pauses and a rather long pause-free interval benefit the attractiveness of a voice.
- The usage of hesitations is reducing the attractiveness. A long average length of hesitations implements a less attractive judgement.
- Single factors can be compensated by others.

In order to investigate a more detailed database a further study is definitely required. Another topic for further study is whether it is possible for male speakers to practice their voices and phonetic behaviour in order to reach a certain state of attractiveness for their voice in the ears of women.

5. COMMENTS

It is conceivable that other factors, such as dialects and sociolects may affect the attractiveness of a voice or even that the results differ depending on the spoken language. This study tried to reduce these effects by using speakers and listeners from the same geographical area and comparable social background. It is planned that this study is repeated with more speakers and listeners and with female voices as well.

6. REFERENCES

[1] Nolan, F. 1983. *The phonetic bases of speaker recognition.* Cambridge , Cambridge University Press.

[2] Hollien, H. 1971. Three major vocal registers: a proposal, *Proceedings of the 7th international Congress of Phonetic Sciences*, Montreal, 320-331.

[3] Dommelen, W.A. van & Moxness, B.H. 1995. Acoustic parameters in speaker height and weight identification: sex-specific behaviour. *Language and Speech*, 38, 267-287.

[4] Bruckert L., Lienard J.S., Lacroix A., Kreutzer M., Leboucher G. 2006. Women use voice parameters to assess men's characteristics, *Proceedings. Biological sciences*, Royal Society of London, Vol.273, Number 1582:83-89.

[5] Koreman, J. 2006. Perceived speech rate: The effects of articulation rate and speaking style in spontaneous speech, *J. Acoust. Soc. Am.* 119, 582-596.

[6] Ní Chasaide, A., Gobl, C. 1997. Voice Source Variation. In: Hardcastle, W., Laver, J. (eds.), *The Handbook of Phonetic Science.* Oxford: Blackwell, 427-461.