

WHAT MAKES A FEMALE VOICE ATTRACTIVE?

Xuan Liu & Yi Xu

Dept. of Speech, Hearing and Phonetic Sciences, Division of Psychology and Language Sciences,
University College London, UK
uclexli@ucl.ac.uk; yi.xu@ucl.ac.uk

ABSTRACT

Vocal attractiveness is highly relevant both for speech technology and for our basic understanding of speech prosody. Yet our knowledge about it is still very limited, and this is particularly true in regard to female voice. The present study explores the nature of female vocal attractiveness by exploring a mechanism that has been shown to be effective in encoding emotions. In a perception experiment, we used utterances produced by a female English speaker with normal, breathy and pressed voice, and acoustically modified them along the size projection dimension which has been previously found to be emotionally relevant. Ten male subjects judged the attractiveness of the utterances, and the results show that a) the most attractive female voice is one that projects a small body size, and b) the most effective acoustic cue for female vocal attractiveness is voice quality.

Keywords: vocal attractiveness, emotion, size projection, evolution, selection pressure

1. INTRODUCTION

What makes an attractive human voice? This question probably fascinates all kinds of people, voice coaches, actors and actresses, public speakers, or those who simply want to improve their personal lives [15, 16]. But equally curious are probably researchers who want to improve the quality of speech synthesis, or to make their robots sound more like human. But our knowledge about voice attractiveness is still quite limited, although much research has been conducted in this area. For example, Zuckerman and colleagues have shown that there is a vocal attractiveness stereotype [15, 16]. Feinberg has found the male voice with low F_0 and smaller formant dispersion to be attractive [4]. What is even less known is what makes an attractive female voice. Fraccaro et al. have found that women raise voice pitch when speaking to men they found attractive [5]. There is anecdotal evidence that a breathy female voice is attractive, but so far there has not been clear experimental

evidence. What is the most lacking, however, is *why* a particular feature makes the voice attractive.

One useful starting point for understanding voice attractiveness is the dimorphism of male and female vocal apparatus and the resulting voice characteristics. It is known that male vocal tract is longer than female vocal tract, that females have higher voice pitch than males, and that female voice is breathier than male voice [6]. It is also known that such dimorphism starts to develop at puberty, i.e., during the time period of sexually maturation. Thus there is little doubt that the dimorphism is sexually motivated. But what is the reason for the specific directions of the dimorphism? Why should the male vocal tract be longer and the female pitch higher? A potential explanation can be found in a link to a theory of animal behavior by Morton, namely, the 'motivational-structural rules' (MS) [9]. According to Morton, animals use body-size related strategies to influence the receivers of the calls. The proposed MS rules influence signal structure in the following ways:

1. Harsh, relatively-low frequency sounds indicate that the sender is likely to attack if further approached or the receiver stays in the same distance.
2. More pure tone like, high frequency sounds suggest that the sender is submissive or appealing if approached or if approaching, or fearful.

According to Morton, the use of harsh and low-pitched voice is to project a large body size so as to scare off the receiver of the signal, while the use of pure-tone like and high-pitch voice is to project a small body size to attract the receiver. Ohala proposed that the same principle, which he referred to as the "frequency code", also applies to human communication [10]. Xu, Kelly & Smillie extends the previous work to a "*bio-informational dimensions theory*", based on the assumption that "human vocal expressions of emotions are evolutionarily shaped to elicit behaviors that may benefit the vocalizer" [14]. These effects would be

achieved by manipulating the vocal signal along a set of bio-informational dimensions: size projection, dynamicity, audibility and association [14].

The size projection dimension is similar to Morton's MS and Ohala's frequency code, which involves three separate acoustic cues: spectral density (which reflects vocal tract length), F_0 and voice quality. These cues can be linked to the male-female dimorphism in that the male voice traits are consistent with cues for aggression and dominance, while the female voice traits are consistent with cues for appeasement and sociability. It would thus be interesting, as the present study will explore, whether what makes a female speaker sound smaller in body size is also what make her voice attractive to male listeners.

2. METHOD

2.1. Stimuli

A female speaker, age 23, with a south-eastern British accent, was asked to produce three emotion-free versions of the sentence "Good luck with your exams", each in one of three voice qualities: normal, breathy and pressed. Here we assume that, with increased spectral tilt from modal voice, and thus reduced high-frequency energy [7], breathy voice is more pure-tone like, whereas pressed voice, with increased high-frequency energy, is harsher and less pure-tone like. The recording was made in a quiet room with a head-mounted condenser microphone (Countryman Isomax hypercardiod). The three recorded utterances were manipulated in terms of three parameters, shown in the first three columns of Table 1. The first two manipulations were directly motivated from what was discussed above. The Final F_0 slope, i.e., the rate of the F_0 fall in the last syllable of the word "exam", was motivated by previous reports that angry or dominant speech may involve steeper F_0 drops than happy or sociable speech [10, 11].

Table 1: Parameter manipulation.

Formant shift ratio	Pitch shift	Final F_0 slope	Voice quality
1.1	+2 st	+15 st/s (height + 3 st)	breathy
1	0	1	normal
0.9	-2 st	-15 st/s (height - 4 st)	pressed

The manipulations were done with a specially written Praat script that applied the "Change gender" function in Praat. The total number of

stimuli were 3 (formant ratio) x 3 (pitch shift) x 3 (F_0 slope) x 3 (voice quality) = 81. The manipulations had no adverse effects on the naturalness of the voice.

2.2. Procedure

Ten young males with an average age of 23 participated as subjects. They were all native speakers of English with no self-reported speech or hearing impairments. They did the listening test individually. The test was run by an Experimental MFC module in Praat on a laptop computer. The task of the subject was to judge the attractiveness of each utterance played through a set of headphones, on a scale of 1-5, with 5 being the most attractive.

3. ANALYSIS

The attractiveness scores were analyzed in a four-way repeated measures ANOVA with Voice quality, Formant shift, Pitch shift and Final slope as independent variables. The results for the four main effects are shown in Table 2.

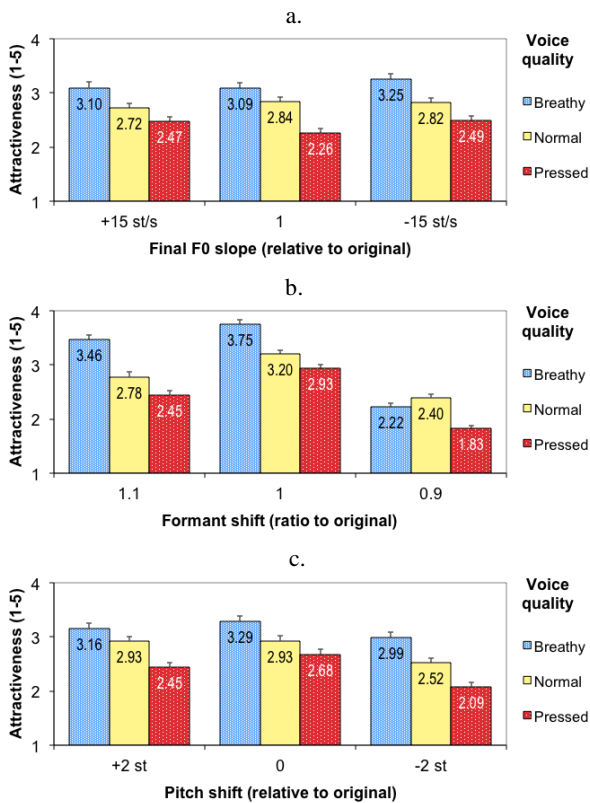
Table 2: Main effects of 4-way repeated measures ANOVA.

Factor	df	F	<i>p</i>
Voice quality	2,18	73.71	< 0.0001
Formant shift	2,18	21.31	< 0.0001
Pitch shift	2,18	11.00	0.0008
Final slope	2,18	2.27	0.1319

As can be seen in Table 2, the main effects of Voice quality, Formant shift and Pitch shift are all highly significant, whereas the effect of Final slope is not significant. There is, however, a significant interaction between Voice quality and Final slope ($F(4,36) = 4.61$, $p = 0.0041$). As can be seen in Fig. 1a, this is probably because the normal slope with pressed voice sounded excessively unattractive.

There is a significant interaction between Voice quality and Formant shift ($F(4,36) = 8.33$, $p < 0.0001$). As can be seen in Fig. 1b, when vocal tract is lengthened (Formant shift ratio = 0.9), the preferences for the three voice quality was different from other vocal tract lengths. But the pressed voice is always the least attractive. A Bonferroni/Dunn post-hoc test shows significant differences between the 0.9 and 1.0 ratio and 0.9 and 1.1 ratio. There was no significant difference between the 1.0 and 1.1 ratios.

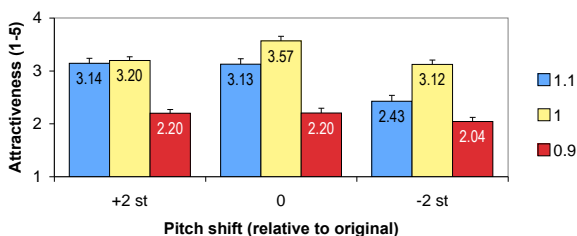
Figure 1: Interactions of Voice quality with Final F_0 slope (a), Formant shift (b) and Pitch shift (c).



The interaction between Voice quality and Pitch shift is only marginally significant ($F(4,36) = 3.00$, $p = 0.0311$). Overall, the original pitch was the most attractive, but there was no significant difference between the original and raised pitch according to a Bonferroni/Dunn post-hoc test. However, the differences between lowered pitch and both original and raised pitch are significant.

Finally, there is a significant interaction between Formant shift and Pitch shift. Fig. 2 shows that, when formants remain neutral or are condensed, the original pitch sounded more attractive.

Figure 2: Interaction of Formant shift and Pitch shift.



4. DISCUSSION

The results of the perception experiment show that voice quality, vocal tract length and pitch height all

significantly affect perceived female vocal attractiveness, but the effect of final F_0 slope is minimum. Of the first three significant factors, voice quality is by far the most important and a breathy voice is always heard as the most attractive. These results are consistent with the hypothesis that vocal attractiveness is closely related to emotional encoding mechanisms and that female attractiveness is mainly achieved by projecting a small body size.

There has been much research on gender-related breathiness in voice. Henton and Bladon reported that breathiness was consistently involved in the production of female speakers. In contrast, much less breathiness was detected in male speakers [6]. More interestingly, they also found that breathiness actually reduced intelligibility. So women seem to employ breathiness in their speech production at a risk to undermine their speech perceptibility.

It has also been reported that there are age related differences in the size of the gap in the rear of the glottis. The anatomy of the larynx has been shown to undergo marked changes from young adulthood to old age [10]. Both the vocal cord vibratory patterns and glottal closure patterns are altered when aging. For example, a high occurrence of glottal gaps has been found in elderly speakers [1]. However, glottal gaps have also been found to occur in young women, and in fact a high incidence of posterior chink is prevalent in this age group [8, 12]. Interestingly, such gaps rarely occur in young men, which suggests the occurrences of glottal gap are unlikely due to age-related inability to close the glottis completely [8]. Rather, it is more likely that young women choose not to fully close their glottis for some functional reasons, i.e. to produce a breathy voice quality. Thus the existence of the posterior chink in female glottis is likely for the sake of producing breathy quality so as to increase vocal attractiveness, as indicated by the present results.

Previous research concerning body size and vocal attractiveness has found that women with large bodies were judged to be low in both facial and vocal attractiveness [3]. This is in line with the current finding that voices with greater spectral density are heard as least attractiveness. However, further shortening from the normal vocal tract length did not generate extra benefit. It is possible that the vocal tract length of our female speaker was already short enough to reach the maximum attractiveness level, and further shortening it might

have made the voice sound like that of a child, which presumably is not generally attractive to men. Further research in this respect is needed.

As seen in Fig. 1b and Fig. 2, the original rather than raised pitch had the highest attractiveness ratings. But there was no statistical difference between the original and raised pitch, and the lowered pitch was always disfavored. Thus the result is still in line with the prediction of the size projection hypothesis, except that there could be some kind of limitation beyond which further pitch raising provides no extra benefit. Again, further research is needed.

Finally, the finding that final F_0 slope had minimal effect on attractiveness suggests that, because such local F_0 contours are more directly relevant for linguistic information, it is less likely to be useful for carrying emotional or attractiveness information. This is despite previous findings suggesting the relevance of F_0 slope for emotional expressions [10, 11].

5. CONCLUSION

The present study has found evidence that female vocal attractiveness is encoded along the same size projection dimension that has been suggested for encoding animal calls and human emotional expressions [9, 10, 14]. That is, an attractive female voice is breathy, with a short vocal tract (though not too short) and high pitch (again, not too high), and all of them serve to project a small body size. Such small-size projection seems to have the effect of attracting the listeners, which is also used in animal calls that express appeasement and submission [9], and human expressions of sociability and happiness [2, 10, 13, 14]. Of the four possible cues for body size projection, we have found voice quality to be the most salient for female vocal attractiveness, and breathy voice to be the most attractive. This seems to point to the functional motivation of the posterior chink frequently found in the female glottis. In general, the current findings demonstrate the effectiveness of the evolutionarily-based approach that has been advocated by Morton and Ohala [9, 10] for a long time but still yet to be widely adopted.

6. REFERENCES

- [1] Biever, D., Bless, D. 1989. Vibratory characteristics of the vocal folds in young adult and geriatric women. *J Voice* 3, 120-131
- [2] Chuenwattanapranithi, S., Xu, Y., Thipakorn, B., Maneewongvatana, S., 2008. Encoding emotions in speech with the size code — A perceptual investigation. *Phonetica* 65, 210-230.
- [3] Collins, S.A. 2000. Men's voice and women's choice. *Animal Behaviour* 60, 773-780.
- [4] Feinberg, D.R., Jones, B.C., Little, A.C., Burt, D.M., Perrett, D.I. 2005. Manipulations of fundamental frequencies influence the attractiveness of human male voices. *Animal Behavior* 69, 561-563.
- [5] Fraccaro, P., Jones, B., Vukovic, J., Smith, F., Watkins, C., Feinberg, D., Little, A., DeBruine, L. In press. Experimental evidence that women speak in a higher voice pitch to men they find attractive. *Journal of Evolutionary Psychology*.
- [6] Henton, C.G., Bladon, R.A.W. 1985. Breathiness in normal female speech: inefficiency versus desirability. *Language & Communication* 5(3), 221-227.
- [7] Klatt, D.H., 1990. Analysis, synthesis, and perception of voice quality variations among female and male talkers. *Journal of the Acoustical Society of America* 87, 820-857.
- [8] Linville, S.E. 1992. Glottal gap configurations in two age groups of women. *Journal of Speech and Hearing Research* 15, 1019-1215.
- [9] Morton, E.S. 1977. On the occurrence and significance of motivational-structural rules in some bird and mammal sounds. *The American Naturalist* 111, 855-869.
- [10] Ohala, J.J. 1984. An ethological perspective on common cross-language utilization of F_0 of voice. *Phonetica* 41, 1-16.
- [11] Scherer, K.R., 2003. Vocal communication of emotion: A review of research paradigms. *Speech Communication* 40, 227-256.
- [12] Sodersten, M., Lindestad, P. 1990. Glottal closure and perceived breathiness during phonation in normally speaking subjects. *Journal of Speech and Hearing Research* 33, 601-611.
- [13] Xu, Y., Kelly, A., 2010. Perception of anger and happiness from resynthesized speech with size-related manipulations. *Proceedings of Speech Prosody 2010*, Chicago.
- [14] Xu, Y., Kelly, A., Smillie, C. Forthcoming. *Emotional Expressions as Communicative Signals*.
- [15] Zuckerman, M., Driver, R. 1989. What sounds beautiful is good: the vocal attractiveness stereotype. *Journal of Nonverbal Behavior* 13, 67-82.
- [16] Zuckerman, M., Hodgins, H., Miyake, K. 1990. The vocal attractiveness stereotype: replication and elaboration. *Journal of Nonverbal Behavior* 14, 97-112.