ANTHROPOLOGIE • XLI/1–2 • pp. 93–98 • 2003

VANESSA ZACHER, CARSTEN NIEMITZ

WHY CAN A SMILE BE HEARD? A NEW HYPOTHESIS ON THE EVOLUTION OF SEXUAL BEHAVIOUR AND VOICE

ABSTRACT: When speaking on the telephone a smiling face can be detected with some certainty by listening to the partner. For an acoustic investigation, 10 subjects were asked to read out 27 words and 4 sentences with and without a smiling face. The utterances were recorded on a Sony DAT-recorder and analysed using WinSAL 1.2a bioacoustic software.

For this test, 141 adult individuals were asked to discriminate among six possible guesses of different facial expressions. The guesses of smiling voices were correct in significantly more cases than random expectation value.

The subsequent analysis showed that smiling has an effect on most acoustic parameters: It raised the energetic maximum of the basal formant as well as the average frequency of the second and the third formants. The amplitude of single words increased. More than 20 other parameters tested were significantly different in smiling and non-smiling voices – included were differences in duration, in the different ways of articulation of vowels and differences between both sexes. These highly complex differences cannot merely be an effect of passive deformations of the vocal tract through smiling. Cerebral control of many unintentional parameters of communication must have a genetic basis. Human courtship behaviour is mostly done in privacy or often in poor light conditions. This may be the reason for many significant differences between the smiling and the non-smiling voice.

KEYWORDS: Smiling voice – Audible – Sex-specific – Human courtship behaviour

INTRODUCTION

Smiling is an universal, innate facial expression of happiness or amusement (Eibl-Eibesfeldt 1972, Dearborn 1900), being used as a social signal (Eibl-Eibesfeldt 1997, Kraut, Johnston 1979) in various positive relations. While facial expressions are predominantly exchanged between two partners, the phylogenetic combination of visual signals with acoustic ones renders this communication more public, being heard by people in the vicinity (cf. the evolution of laughter: Niemitz *et al.* 2000). Everyday experience shows that it is, at least sometimes, possible to identify aurally the emotional state of a speaker, solely by listening to his or her voice, for example in a phone call.

This is caused by a side effect of mimic changes of smiling, namely an alteration of the shape of the vocal tract. The mouth orifice widens, the angles of the mouth drawn backwards shorten and enlarge the vocal tract (Shor 1978). Also, a smiling voice is almost never loud, since it is difficult to raise one's voice while smiling. Thus, a smiling speech is a short distance communication, among friendly social partners, among close friends and between sexual partners. For all these reasons, we investigated how certainly a smile can be discerned by listening to a speaker's voice. Moreover, we searched for the differences of acoustic parameters in neutral and smiling speech, which might be responsible for identification of smiling.

TABLE 1. List of the 27 words and 4 sentences used in the experiment, in orthographical spelling and in SAMPA-notation (see text).

Wort	SAMPA	Wort	SAMPA
Ried	[ri:t]	Füllt	[fylt]
Politik	[poli´ti:k]	Höhle	[´h2:l@]
Beet	[b´e:t]	Ökonom	[2ko´no:m]
Bett	[bEt]	Göttlich	[g9tlIç]
Hase	[´ha:z@]	Ober	[´o:b6]
wähle	[´vE:l@]	Weit	[vaIt]
Dach	[dax]	Frau	[fraU]
Ja	[ja:]	Neu	[nOY]
Post	[pOst]	Uhr	[u:6unsilbisch]
Hut	[hu:t]	Lang	[laŋ]
Musik	[mu´sik]	Genie	[Ze´ni:]
Muss	[mUs]	Schal	[Sa:l]
Metall	[me´tal]	Ich	[Iç]
Hüte	[hy:t@]		
Könnten Sie bitte die öffentliche Rufnummer wählen?			
[´k9nt@n zi: bit@ di: 9´f@ntliç@ ru:fnum6 vE:l@n]			
Hattest Du eine angenehme Fahrt ?			
[hat@st du: aIn@ ang@ne:m@ fa:rt]			
Die besten Wünsche für Deine Frau.			
[di: bEst@n vynS@ fy:r daIn@ fraU]			
Ich bin erfreut, von Ihnen zu hören.			
[Iç bin ErfrOYt fOn i:n@n tsu: h2:r@n]			

METHODS

To examine the effect of smiling in the voice, 10 test subjects (5 males and 5 females) read a list of 27 words and 4 sentences (*Table 1*) in two different speech types, first normally, i.e. without any emotional facial expression, and second, with a smiling face. They were instructed by standardized information not to imitate the emotional state, but to form the facial expression. The speech was recorded using a Sony digital audio tape-recorder with a frequency range of 30–70.000 Hz.

In a first step we made a discrimination analysis of the smiling and the normal voice. Therefore, six identification tapes were prepared from the words and the sentences spoken by the speakers, using the computerized speech-editing software system WinSAL 1.2a. Each tape was of 14–30 trials of one sentence, or a group of 9 words in random order. Several groups of students were asked to select 1 out of 6 given facial expressions of the speaker in each trial, such as: anxious; furious; normal/neutral; sad; smiling; or others. Thus, the guessing probability was 16.7%.

In a speech analysis of the utterances of the ten speakers we investigated what parameters are responsible for the audible recognition of smiling. Therefore we measured the following parameters: 1. the median of the frequency [Hz] of the maximum of the basal frequency (F_{0max})of vowels; 2.–4. the median of the frequency [Hz] of the first three formants (F_1, F_2, F_3) of vowels; 5.–6. the median of the relative amplitude [dB] of words and sentences and the median of the duration [msec] of 7. words, 8. vowels and 9. sentences.

For the measurements of vowels we chose one example of each phonetic vowel of the 27 words. For these vowels, we also analysed the differences between different articulation modes as used in linguistic literature. In order to find differences between the two speech types, the median of the smiled type minus the median of the neutral/normal type were used for calculation. Significant values were obtained applying Wilcoxon-test of ranked pairs, the significant thresholds being $p<.05$; $p<.01$; $p<.001$. The Wilcoxon-test was also used for the calculation of sex-specific differences. The quartiles are not presented here for technical reasons.

RESULTS

Audible recognition of smiling speech

In 27.1% of all cases (N=1,736) the listening subjects recognized the utterances with a smiling facial expression as smiling. The normal/neutral utterances were recognized correctly in 42.5% of all cases (N=1,665). While the expected random probability for correct answers was 16.7%, both percentages were far above expectation value and highly significant (both $p<.001$).

Speech analysis

The speech analysis shows that smiling has an effect on most acoustic parameters investigated here (*Figure 1*). The

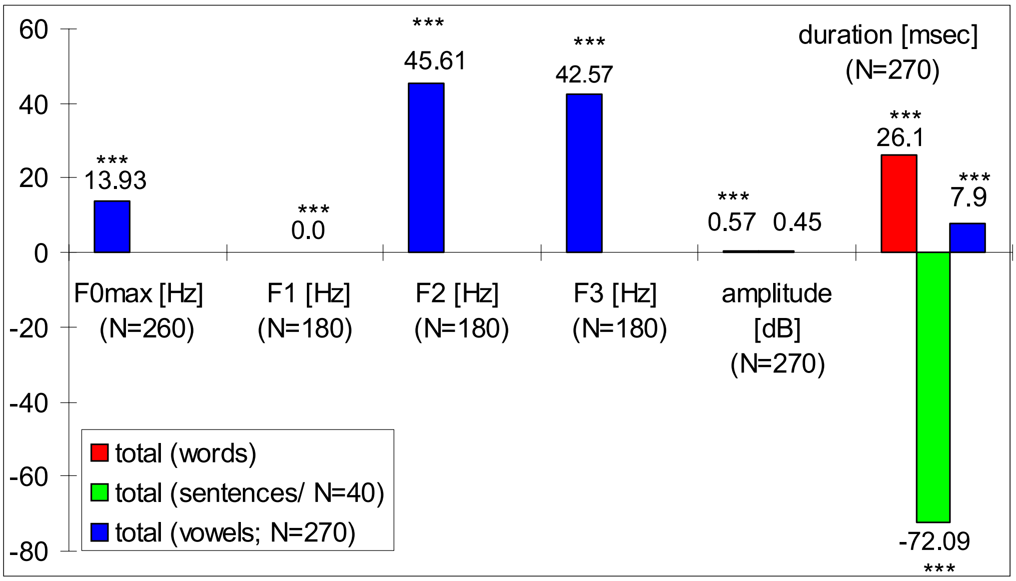

FIGURE 1. Differences represent the average value for the smiling speech minus the neutral/normal one. Significance values were obtained applying Wilcoxon ranked pairs test: *p<.05; **p<.01; ***p<.001. N = the respective number of rating values.

frequency of the energetic maximum of the basal formant F_{0max} increased from the normal speech to the smiling one by ΔF=13.9 Hz (N=260, p<.001). For the second formant this increase measured ΔF=45.6 Hz (N=180, p<.001). For the third formant, an increase of 42.6 Hz (N=180, p<.001) was found on the average of all subjects. Only the average frequency of the first formant between smiling and non-smiling speech was not significant; this may depend on measuring problems. The relative amplitude of single words increased by 0.57 dB (N=270, p<.001).

On the average, a smiling face while speaking extended the duration of single words by 26 msec (N=270; p<.001). On the average, the increase of the duration of vowels while smiling was 7.9 msec (N=410; p<.001). But interestingly enough, the average duration of sentences decreased in smiling utterances by –73 msec (N=40; p<.01) by reducing the duration of other elements of speech (see Discussion).

The effect of smiling on the different ways of articulation of various vowels yielded interesting differences for several acoustic parameters investigated. In smiling faces, while talking, the duration of most vowels showed an effect dependent to the position of the tongue. Only the vowels with deep tongue position ([a], [a:]) showed no difference for smiling speech (*Figure 2*), while vowels with a middle,

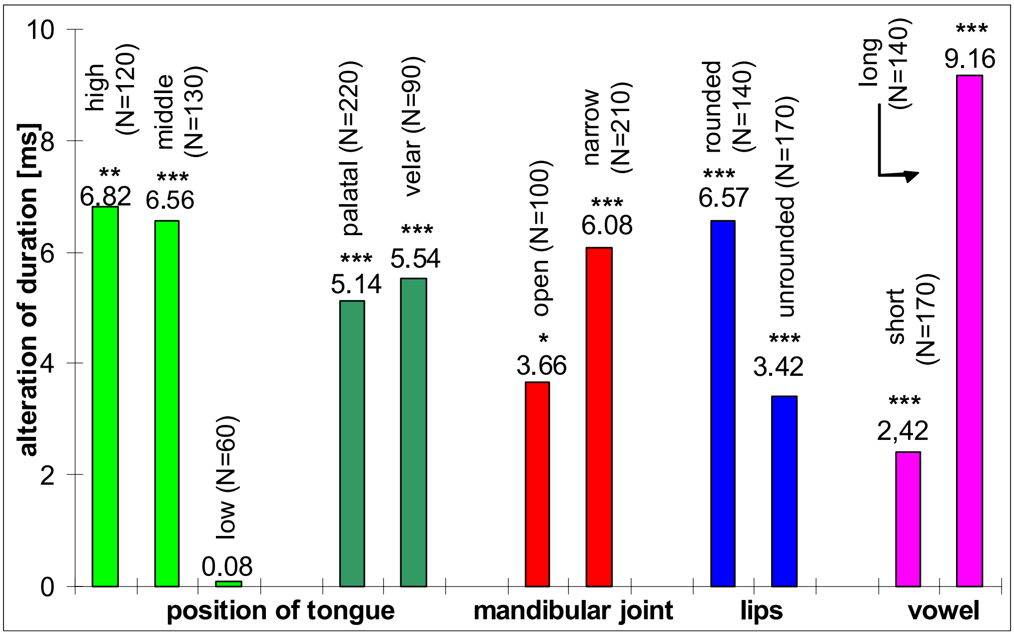

FIGURE 2. Differences of duration represent the average value for the smiling speech minus the neutral/normal one. Significance values were obtained applying Wilcoxon ranked pairs test: *p<.05; **p<.01; ***p<.001. N = the respective number of rating values.

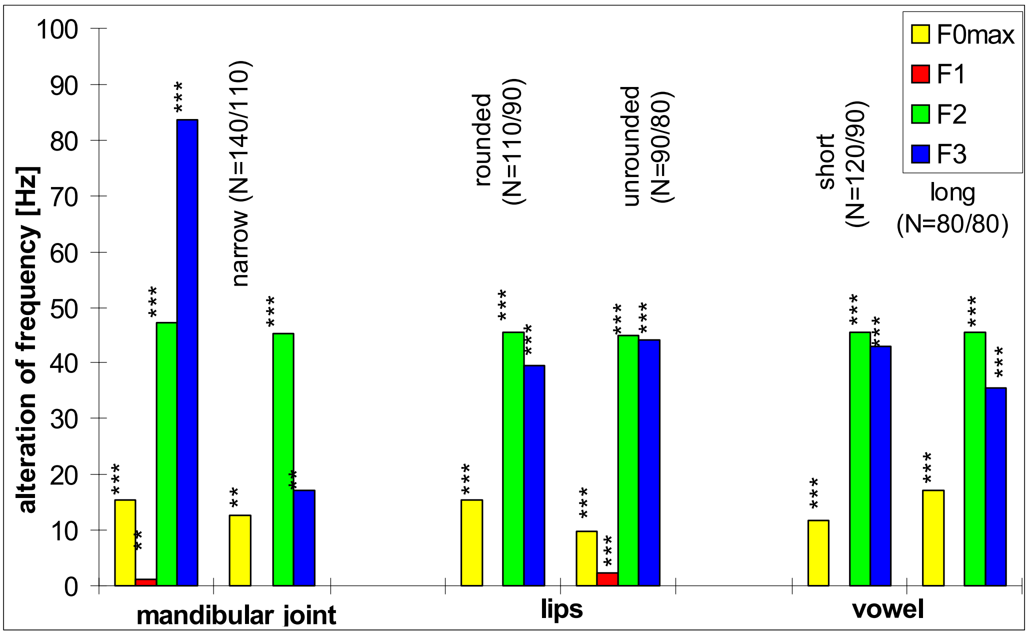

FIGURE 3. Differences of basal frequency (F_{0max}) and the formants (F_1, F_2, F_3) represent the average value for the smiling speech minus the neutral/normal one. Significance values were obtained applying Wilcoxon ranked pairs test: *p<.05; **p<.01; ***p<.001. N = the respective number of rating values.

high, velar or palatal position increased in their average duration. For closed vowels (i.e. with a narrow angle between the tooth rows; a minor opening of the mouth; i.e.[I], [y], [u]) and for rounded vowels (i.e. [u], [o]) , this increase in duration while smiling was twice as high as for those with a wider angle between the tooth rows (open; i.e. [E], [a]) and for those without a round orifice (unrounded; i.e.[I], [a], [e]). Long vowels while smiling had the highest increase of duration, whereas short vowels showed to have the lowest increase.

While smiling, all modes of tongue position in vowels were linked to an increase in the maximum of the basal frequency (F_{0max}), (*Figures 3* and *4*). This increase was twice as high in vowels with a high (ΔF=19.8 Hz, N=80, p<.001; i.e. [I], [y], [u]) and velar (ΔF=18.9 Hz, N=60, p<.001; i.e.[O], [9], [e:]) position of the tongue in comparison to middle (ΔF=9.0 Hz, N=90, p<.01; i.e. [o], [e], [2:]) and palatal (9.6 Hz, N=20, p<.001; i.e. [I], [e], [a]) positions. Also we found differences within the other groups of articulation, all of which reached a significance level of

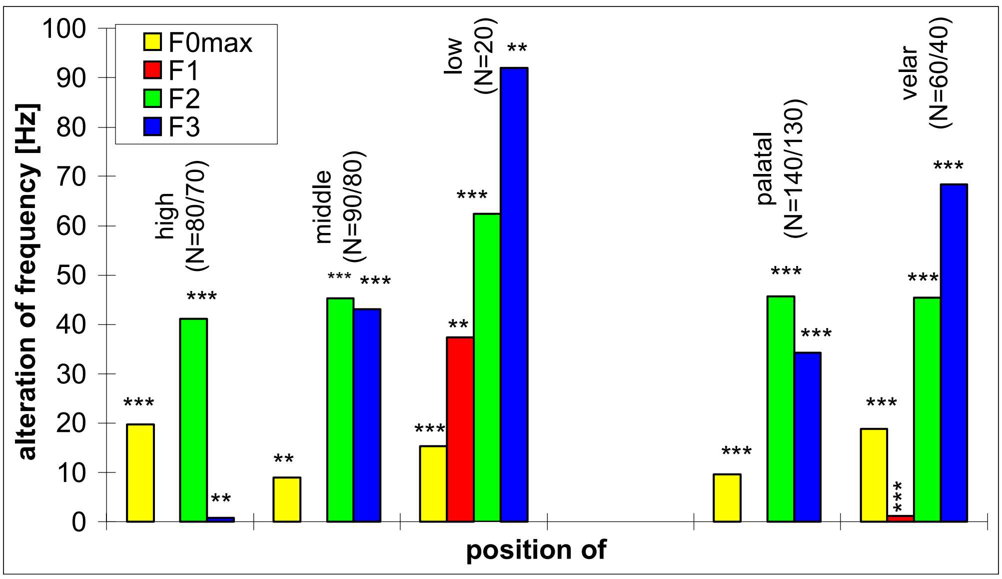

FIGURE 4. Differences of basal frequency (F_{0max}) and the formants (F_1, F_2, F_3) represent the average value for the smiling speech minus the neutral/normal one. Significance values were obtained applying Wilcoxon ranked pairs test: *p<.05; **p<.01; ***p<.001. N = the respective number of rating values.

TABLE 2. Sex-specific differences between smiling and normal facial expression in Percent of the Median. Significance values were obtained applying Wilcoxon ranked pairs test: *p<.05; **p<.01; ***p<.001. N = the respective number of rating values.

Difference between	Women	Higher modification in	Men	Difference between sexes [%]	Significance
Duration of words	29.9 ms	>	24,4 ms	**23%**	*
Duration of vowels	9.4 ms	>	6.6 ms	**43%**	*
Duration of sentences	–105.8 ms	>	–35.4 ms	**199%**	**
Amplitude of words	0.56 dB	=	0.58 dB	**4%**	**n.s.**
Amplitude of sentences	0,09 dB	<	0.55 dB	**511%**	
$\mathbf{F_{0max}}$	18.3 Hz	>	12.78 Hz	**43%**	**
$\mathbf{F_1}$	(14.33 Hz)	?	(0 Hz)	**?**	**n.s.**
$\mathbf{F_2}$	45.6 Hz	>	44.3 Hz	**3%**	**
$\mathbf{F_3}$	40.6 Hz	<	46.8 Hz	**15%**	**

p<.001: The average increase in the maximum of the basal frequency (F0max) was higher for open vowels (i.e. with a wider angle between the tooth rows) (ΔF=15.6 Hz, N=60) than for vowels with nearly closed jaws (ΔF=12.6 Hz, N=140), and also for the vowels with rounded lips (ΔF=15.5 Hz, N=110) in comparison to vowels without rounded lips (ΔF=9.6 Hz, N=90), as well as for the long vowels (ΔF=17.3 Hz, N=80) in relation to short ones (ΔF=11.7 Hz, N=120).

We observed very different effects of smiling on the three formants of spoken vowels when regarding the position of the tongue. But whenever there was a difference in the frequency between neutral and smiling speech, the frequency was higher in the smiling one. The vowels spoken with a low tongue position showed the highest increase in frequency (for ΔF, N and p-values see: *Figure 3*). F_2 and F_3 reached, most constantly, the highest levels of significance. Moreover, there was a tendency for the velar vowels to be submitted to a higher increase of F_{0max}, F_1 and F_3 than palatal ones, whereas the increase value for F_2 was almost the same in both types of vowels.

With respect to the tooth rows, the shape of the lips, and the duration of vowels, we found that the increase of frequencies were much less dependent from such differences of articulation, although 17 out of 18 such comparisons of F_{0max}, F_2 and F_3 in smiling and non-smiling vowels reached the highest level of significance (*Figure 4*).

Sex-specific differences in smiling speech

In all parameters there is the same kind of effect of smiling on speech in both sexes. Sex-specific aspects are just related to the level of differences between a smiling and a normal face (*Table 2*).

The average increase in duration of single words, spoken was 24.4 msec in smiling men, but significantly higher in smiling women: 29.29 (N=135 each; p<.001). Thus, the increase of duration of vowels was 43% higher in female than in male test subjects, while the lengthening of words was 23% higher in women than in men. Also the shortening of duration of sentences was about 3 times greater in our female subjects than in the male ones (men: –35.4 msec; n.s.; women: –105.8 msec; p<.001; N=20 each). The average increase of the relative amplitude of words was equal in both sexes, while the relative amplitude of sentences was 511% higher in men (0.55 dB; N=20; p<.01) than in women (0.09 dB; N=20; n.s.).The average increase of the maximum of the basal frequency F_{0max} in men amounted to 12.8 Hz (N=130; p<.001). For women, however, it was much higher and reached 18.3 Hz (N=130; p<.001). The increase of the second formant while smiling seems to be equal, in men it increased by 44.3 Hz (N=90; p<.001) and by 45.6 Hz in women (N=90; p<.001). But this difference of only 3% percent between both sexes is significant. In smiling men, the increase of the third formant was 15% higher in men than in women. This demonstrates that the effects of smiling on speech are different in men and women.

DISCUSSION

The experiment demonstrates that smiling can be recognized aurally. So, it must have perceptible effects on speech. An increase of duration of words and vowels, a shortening of the duration of sentences, an increase of the frequency of the maximum of basal frequency and also of the second and third formants, as well as an increase of the relative amplitude form part of these effects of smiling on speech.

The changes that smiling causes on the duration of words or sentences seem contradictory, since the duration of vowels and words increases, while there is a decrease of duration in sentences. But these opposed effects depend on the fact that the rate of speaking is lower for vowels than for consonants and also lower for stressed syllables (Bergmann *et al.* 1988). Also, Burkhardt and Sendlmeier (1999) have shown that, when human speech is simulated, a lengthening of stressed syllables of about 24% and a shortening of unstressed syllables of about 20% yield a high rating for the emotion "joy". There are typically differences in stress between words and sentences. Single words have more stressed syllables and fewer syllables without stress than words. The opposed effect of shortening of unstressed syllables causes the shortening of sentences.

In our experiment we found that, in smiling, closed vowels have a higher increase in duration than open ones. This may depend on the fact that smiling prevents the mandibular joint from closing and thus, the anatomical distances for forming of vowels may be longer while smiling. This effect may also be the cause for the higher increase for vowels with rounded lips.

The increase of the relative amplitude through smiling may depend on the width of the mouth orifice. Another explanation could be the coupling between the tension of the face and larynx musculature (Tartter 1980), as an increase in tension of larynx also generates an increase in amplitude (Murray, Arnott 1993). The increase of the maximum of the basal frequencies while smiling is a result of the shortening of the vocal tract. Additionally also in this case, an increase of the tension of the laryngopharynx (Tartter 1980) and a change of the position and the shape of larynx (Riordan 1977) may all be responsible for our observation, each in a different way.

The increase of frequency of the formants while smiling depends, at least partly, on the shortening of the vocal tract as well (Laver 1991). Moreover, there is an increased rigidity of the mouth cavity (Laver 1980) caused by muscle tonus. It may also depend on the opening of the mandibular joint angle (Lindblom, Sundberg 1971). The effect of the opening of the angle between the jaws shows that the increase is higher for open vowels than for more closed ones. The high increase for velar and deep vowels may depend on the tonus of the mouth cavity, because the oral space is not reduced in these vowels.

There is an enormous array of highly significant phonetic and acoustic effects of smiling on speech. They cannot be caused merely by the shortening of the vocal tract or by other simple side-effects of the muscle actions for the facial expression. There are significant differences in mostly all parameters measured for words, vowels and sentences. Last but not least in a number of respects, a smiling voice is different from a non-smiling voice in men and women, i.e. both sexes do not only possess dimorphic voices, but furthermore, when smiling they change their voices differently.

From all this, we conclude that these many differences cannot be controlled by consciousness as they are uttered without the speakers' knowledge. Therefore, they must be determined genetically. Like facial expression, the smiling voice seems to be an inborn acoustic friendly positive signal.

However, inborn social behaviour of a species can only have become a feature of this taxon by means of natural selection. But of which kind are the selective factors?

The perception of a voice as a smiling one is certainly more important in twilight or darkness than at bright light. At the same time, smiling is an essential part of courtship behaviour in man. We therefore hypothesize that a discernable smiling voice at twilight or in the dark may replace or supplement the facial expression. Add to this, the facial expression is, in important situations, sometimes hardly or even not visible. Sexually dimorphic smiling voices may further enhance this function in human courtship behaviour.

REFERENCES

BERGMANN G., GOLDBECK T., SCHERER K. R., 1988: Emotionale Eindruckswirkung von prosodischen Sprechmerkmalen. *Zeitschrift für experimentelle und angewandte Psychologie* XXV, 2:167–200.

BURKHARDT F., SENDMEIER W. F., 1999: Simulation emotionaler Sprechweise mit konkatenierender Sprachsynthese. *Vortrag, 30. Jahrestagung der GAL 1999 in Frankfurt Main.*

DEARBORN G. V. N., 1900: The nature of smile and laugh. *Science* XI: 851–856.

EIBL-EIBESFELDT I., 1972: Similarities and differences between cultures in expressive movements. In: R. A. Hinde (Ed.): *Nonverbal Communication*. Cambrigde University Press, Cambridge.

EIBL-EIBESFELDT I., 1997: *Die Biologie des menschlichen Verhaltens*; 3. Edition Seehamer Verlag GmbH, Weyarn.

KRAUT R. E., JOHNSTON R. E., 1979: Social and emotional message of smiling: an ethological approach. *J. of Personality Social Psychology* XXXVII: 138–166.

LAVER J., 1980: *The phonetic description of voice quality*. Cambridge University Press, Cambridge. 141–156 pp.

LAVER J., 1991: *The Gift of Speech*. Edinburgh University Press, Edinburgh. 217–218 pp.

LINDBLOM B. E. F., SUNDBERG J., 1971: acoustical consequences of lip, tongue, jaw, and larynxmovement. *J. of the Acoustical Society of America* L: 1166–1179.

MURRAY I. R., ARNOTT J. L., 1993: Towards the simulation of emotion in synthetic speech: a review of literature on human vocal emotion. *J. of The Acoustical Society of America* XCIII, 2: 1097–1108.

NIEMITZ C., LOI M., LANDERER S., 2000: Investigation on human laughter and its implications for the evolution of the hominid visual communication. *Homo* LI: 1–18.

RIORDAN C. J., 1977: Control of vocal-tract length in speech. *J. of the Acoustical Society of America* LXII, 4.

SHOR R. E., 1978: The production and judgement of smile magnitude. *J. of General Psychology* XCVIII: 79–96.

TARTTER, V. C., 1980: Happy talk: perceptual and acoustic effects of smiling on speech. *Perception & Psychophysics* XXVII: 24–27.

Vanessa Zacher
Carsten Niemitz
Department of Biology
AG Humanbiology
Freie Universität Berlin
Albrecht-Thaer-Weg 6
14195 Berlin, Germany
E-mail: zacher@zedat.fu-berlin.de
cniemitz@zedat.fu-berlin.de