

FORMANT CHARACTERISTICS OF HUMAN LAUGHTER

D.P. Szameitat (1, 2), C.J. Darwin (3), A.J. Szameitat (2), D. Wildgruber (1), A. Sterr (2),
S. Dietrich (1), K. Alter (4)

Universität Tübingen, Germany (1), University of Surrey, UK (2), University of Sussex, UK (3),
University of Newcastle, UK (4)

d.szameitat@surrey.ac.uk, cjd@sussex.ac.uk, a.szameitat@surrey.ac.uk, dirk.wildgruber@med.uni-tuebingen.de,
a.sterr@surrey.ac.uk, dietrich@cbs.mpg.de, kai.alter@newcastle.ac.uk

ABSTRACT

Although laughter is an important aspect of non-verbal vocalization, its acoustic properties are still not fully understood. Here we provide new data on the spectral properties of laughter. We measured fundamental frequency and formant frequencies of the vowels produced in laughter syllables. In accordance with theoretical predictions and prior observations laughter was mainly based on central vowels. Furthermore, laughter syllables showed higher formant frequencies than normal speech vowels; in particular F1 values could be as high as 1300 Hz for male speakers and 1500 Hz for female speakers. These exceptionally high F1 values might be based on the extreme positions adopted by the vocal tract during laughter in combination with physiological constraints accompanying production of a “pressed” voice.

Keywords: laughter, formant, vowel, nonverbal, F1.

1. INTRODUCTION

The acoustical signal of laughter has unique structural features. It consists of a series of repeated syllables produced on a staccato outward breath. Each syllable typically consists of a fricative (aspirated “h” sound) followed by a vowel element [25]. Moreover, laughter can be produced with extreme voice characteristics (e.g. squealing), with its pitch being up to 1245 Hz for male speakers and 2083 Hz for female speakers, respectively [1]. During production of such sound utterances the vocal tract can be under great physiological strain. Furthermore, during laughter the mouth can be opened very wide. This extreme articulation is likely to produce extreme acoustic consequences, such as very high F1 frequencies.

The most extensive study of the spectral properties of laughter was done by Bachorowski and colleagues [1]. However, although females should

have higher formant frequencies than males because of their shorter vocal tract length [20], for some of the formants (i.e. F4 & F5) Bachorowski et al.’s outcomes [1] were not in line with this prediction. Since the authors themselves suggested that this result might be due to peculiarities of the analysis performed, there is a need for further analyses. Other studies that have investigated spectral properties of laughter examined either only a small number of subjects [3] or analysed only two formants [16].

Our study measured the fundamental frequency and the frequency of the first five formants of vowels in laughter syllables produced in various emotional contexts. We also determined vowel elements by comparing F1-F2 plots with Hillenbrand et al.’s speech vowel representation [11].

2. METHODS

2.1. Sound recordings

Eight professional actors (3 male/ 5 female) produced laughter in various emotional contexts (joy, tickle, *schadenfreude* [to laugh at another’s misfortune], sneering). Recordings took place in a sound proof booth, using a DAT recorder (TASCAM DA-P) with a speaker-microphone (Sanyo MP-101) distance of circa 0.5 m. Recordings were digitized (16 bit / 48 kHz), normalized, and cut into individual laughter sequences.

2.2. Sound material

We excluded laughter sequences that contained words, interjections, or background noise, or were of short duration (< 3s) or low amplitude (with non-detectable pitch).

The stimulus set consisted of 125 laughter sequences (49 male) with 10-22 sequences per speaker. Formant frequency measurements were

obtained for 3932 laughter syllables (1689 male / 2243 female).

2.3. Acoustical analysis

Extraction of mean fundamental frequency (F0) and mean frequency of five formants (F1-F5) of each laughter syllable was conducted in Praat 4.02.04 [6]. Fundamental frequency analysis was based on an acoustic periodicity detection using an accurate autocorrelation method [5]. This method allows reliable pitch extraction also for vocalizations which are not fully voiced. Maximum pitch search range was determined by visual inspection, by overlaying the automatically extracted pitch contours with a narrowband FFT-based spectrogram (30 ms, Gaussian window, pre-emphasis +6 dB/octave). Formants were extracted performing a short-term spectral analysis (Gaussian-like window, LPC analysis, Burg algorithm, see [7, 21]), approximating the spectrum of each analysis frame by five formants. Ceiling of the formant search range was 5000 Hz for male and 5500 Hz for female speakers, respectively.

Laughter sequences were segmented in the time domain according to individual laughter syllables (burst of energy of (un)voiced exhaled breath having a single vocal peak). Boundaries of a syllable were determined visually in the amplitude-time spectrum (distinct rise of energy from background noise into a single vocal peak). For syllables with ambiguous outcome in the automatic formant extraction, formant-peak locations were examined by visual inspection on a random basis. For this, the automatically detected formant bands were overlaid with a broadband FFT-based spectrogram (5 ms, Gaussian window, pre-emphasis +6 dB/octave). Formant measurements were not taken from laughter syllables which were unvoiced, produced with closed mouth, or where spectral measurement extraction was uncertain.

To determine vowel quality of the laughter syllables, F1-F2 plots were calculated for each individual speaker and mapped with the speech vowel representation according to Hillenbrand et al. [11].

3. RESULTS

Table 1 shows average fundamental frequency and formant frequency measurements for laughter syllables produced by male (1689 syllables) and female (2243 syllables) speakers. Statistical tests revealed that in all six acoustical parameters female speakers had higher frequency values than

male speakers (independent-samples t-tests, $t(6) = 2.657 - 5.027$, all $p < .05$, Bonferroni-corrected).

Table 1: Frequency measurements for fundamental frequency (F0) and first five formants (F1-F5) for male and female speakers. s.d. standard deviation.

[Hz]	Females	s.d.	Males	s.d.
<i>F0</i>	476	107	199	8
<i>F1</i>	924	128	728	11
<i>F2</i>	1699	93	1530	71
<i>F3</i>	2995	89	2700	58
<i>F4</i>	3842	152	3472	179
<i>F5</i>	4600	117	4184	264

Surprisingly, in 26% of all vowel elements in laughter syllables F1 frequencies were higher than 1000 Hz ($n=1021$), with male speakers showing maximal values up to 1300 Hz and female speakers up to 1500 Hz. Thus, first formants of several laughter syllables had exceptionally high values in comparison with speech vowels [e.g. 11, 20]. These syllables very often sounded as though they had been produced with a hard or “pressed” voice.

According to Hillenbrand et al.’s [11] standard-vowel-space-representation vowel elements of female speakers fall mainly into the (Λ) and (ɑ) range, with some vowel elements falling in the (ɪ), (e), (æ), (ɛ), (ɜ), (c) and (U) range (Fig. 1). Vowel elements of male speakers fall mainly into the (ɜ), (Λ) and (ɑ) range, with some vowel elements falling into the (i), (ɪ), (e), (æ), (ɛ), (c), (U) and (o) range (Fig. 2).

Analysis on the basis of individual male speakers revealed that all but c.10 of the vowel elements falling into the (i), (ɪ), (e) and (æ) range had been produced by the same male speaker. Thus, laughter syllables were predominantly based on central vowels, with vowel height varying from mid (ə) to open (a), probably because of changes in jaw opening.

Analysis of vowel quality according to speaker identity revealed that differences between vowel elements are based mainly on speaker identity. In other words, individual speakers tend to use a constant set of vowel elements (low intra-personal variability).

Figure 1: F1-F2 plot for female speakers with vowel representation according to Hillenbrand et al. [11].

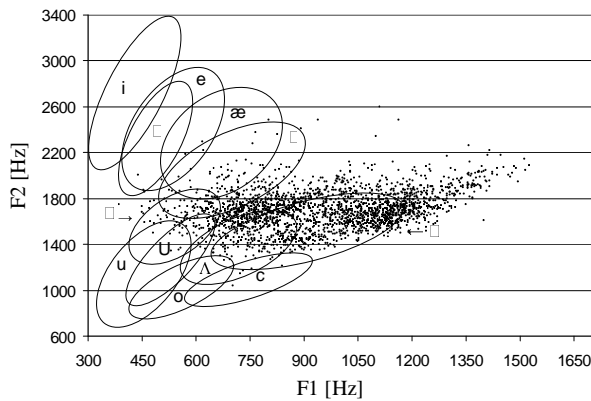
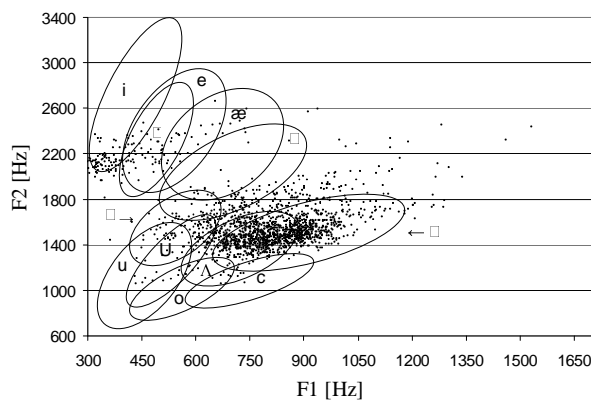


Figure 2: F1-F2 plot for male speakers with vowel representation according to Hillenbrand et al. [11].



4. DISCUSSION

4.1. Fundamental frequency

The mean F0 was 199 Hz for male and 476 Hz for female speakers, respectively, which is well within the range of previously reported F0 for laughter (mean-F0 range males (females) 160-502 Hz (126-424 Hz) [1, 3, 4, 14, 16, 17, 19, 22, 23]). Thus, our data are in accordance with the finding that fundamental frequency in laughter is higher than in speech [17-19, 23, 24].

4.2. Formant frequencies

While frequency measurements of the second to fifth formants fell within the range previously reported for laughter [1, 16] the first formant showed much higher frequencies than expected (here: 924 Hz (females), 728 Hz (males); Bachorowski et al. 2001: 653 Hz / 535 Hz).

High F1 values could be due to erroneous formant extraction. For instance, in high pitched sounds harmonics are widely spaced, so that the fundamental frequency can be higher than the actually articulated F1. The second formant may

then be measured as F1. However, this artefact is unlikely to be the reason that we have obtained such high F1 values. Our high-F1 syllables were not particularly high pitched, but were characterised by a wide range of F0 values (for all F1 > 1000 Hz, females: range 81-1486 Hz, n= 878; males: range 155-404 Hz, n=50). In addition, visual inspection of broadband spectrograms of a random selection of very high-F1 syllables showed sufficient energy in lower frequency bands for an actually lower F1 to have been revealed. Finally, if the true F1 had been missed and F2 consequently identified as F1, then all following formants should be much higher as well (Paul Boersma, personal communication), but this was clearly not the case. Alternatively, so-called pseudo formants (reflecting turbulences in the air flow) may account for the high F1 values. However, pseudo formants are characterised by a high formant bandwidth (>500 Hz [13]), which we observed only in 3.5% of the high-F1 syllables. In addition, almost all examined syllables showed clear harmonic structure. Taken together, it seems very unlikely that the high F1 values are caused by erroneous analysis.

Another cause of the high F1 values may be found in physiological changes in the vocal tract. Firstly, lowering the jaw results in a raised F1 [26]. For instance, soprano singers can raise their F1 up to approx. 1050 Hz (to tune it to their F0) by opening the jaw very wide [12]. Secondly, certain voice qualities associated with narrowing the pharynx lead to a raised F1 [15]. Remarkably, most of the high F1 syllables were produced with a “pressed” voice which may well stem from physiological constraints in the pharyngeal region, such as a lower pharyngeal constriction. Therefore, it seems likely that the currently observed high F1 values are the result of a combination of wide jaw opening and pharyngeal constriction.

A possible explanation why other studies have not yet identified such high F1 frequencies for human laughter is that they may have used laughter which was less expressive, i.e. laughers’ arousal may have been lower than for our material. For instance, in the study of Bachorowski et al. [1] subjects laughed while watching funny video clips in an experimental setting, partly being together with strangers (for a similar approach see [16]). These circumstances may have inhibited the subjects’ laughter response. This inhibition may have led to less extreme articulation, and conse-

quently a lower F1. In contrast, in our study actors were asked to put themselves fully into the emotional contexts given in the instructions, so that laughter might have been produced more expressively.

Another reason might be that in the current study the stimulus set was based on laughter produced by actors, and therefore might differ in its acoustical properties in comparison to spontaneously emitted laughter. Exhaustive acoustical analysis (not reported) revealed that the acoustical properties of the recorded laughter of our study showed no fundamental differences to recent findings for spontaneously emitted laughter [27] (for a similar finding see [2]). The only exception was the longer duration of the laughter, which was introduced by explicit instruction given to the actors since for further planned studies laughter sounds with longer duration were needed. Empirical tests, investigating if people can tell the difference between spontaneously emitted laughter and laughter produced by actors could give new insights on how representative the latter is of human laughter, in general.

A final explanation is that laughter was recorded in a variety of different emotional contexts, which leads to the fact that laughter was emitted with a variety of different voice characteristics [27].

4.3. Differences in speaker sex

Fundamental frequency was higher in female than in male speakers (cf. [22, 23]) with F0 being up to 1765 Hz for female speakers and 595 Hz for male speakers, respectively ([1]: males (females) 1245 Hz (2083 Hz); see also [18] for children: 3200 Hz).

For all five formants females had higher average frequencies than males, which is in accordance with females having a shorter vocal tract than males [20]. Therefore, the current data contradict some of the previously reported findings [1, 16] which found for some of the formants either no differences, or even higher frequencies for males than for females.

4.4. Vowels

Regarding the mapping of the F1-F2-plots for laughter with the speech vowel representation according to Hillenbrand et al. [11] it should be noted that both data sets consist of different speakers, hence different vocal tract lengths.

Therefore, the direct comparison may be prone to some misidentification of vowels. However, outcomes of IPA transcription confirmed our results that mainly central vowels are produced in laughter.

The finding that our laughter consisted predominantly of central sounds is in line with the general hypothesis of Ruch and Ekman [25] and other recent data [1, 22]. The use of central vowels is in accordance with physiological constraints accompanying production of laughter: the vocal tract is in a relaxed position, moreover, raised lip corners and wide jaw opening leave little room for articulation [24]. However, some of our laughter syllables were non-central sounds, as also reflected in previous work [3, 8, 25]. The reason for the production of non-central vowels is not fully understood. Ruch suggested that non-central vowels may be indicators of different emotional qualities underlying the laughter [24]. However, recent findings [27] are not in line with this prediction for the emotional connotations of the laughter investigated in the present study. Furthermore, non-central vowels are also produced when people laugh in a single behavioural context [1], therefore variability in the emotional or situational context seems not to be the leading factor for the production of non-central vowels. Alternatively, use of non-central vowels might be related to intra-individual differences. Previously, it was speculated that each person has their own characteristic laughter sound [8, 10, 18]. This hypothesis is supported by our data, as we found that individual speakers tended to use a constant set of vowel elements, but inter-individual variability was high. To fully understand the use of non-central vowels further investigation is needed.

5. CONCLUSION

In conclusion, these findings indicate that (i) laughter syllables are predominantly formed with central vowels, although others can occur; (ii) formant frequencies show typical gender effects with higher frequencies in female speakers; (iii) compared to speech production, the first formant of laughter vowels is occasionally characterized by exceptionally high frequencies which may be the result of a wide jaw opening and/or pharyngeal changes in "pressed" voice; (iv) the vowel elements during laughter showed a relatively stable individual pattern, whereas the between-subject variability was considerably higher.

6. ACKNOWLEDGEMENTS

We would like to thank the Marie Curie Foundation and the German Research Foundation (Deutsche Forschungsgemeinschaft) for financial support (D.P. Szameitat), DFG AL 357/1-2 (K. Alter) and DFG WI 2101/2 (D. Wildgruber).

7. REFERENCES

- [1] Bachorowski, J.-A., M.J. Smoski, and M.J. Owren. 2001. The acoustic features of human laughter. *Journal of the Acoustical Society of America*, 110(3), 1581-1597.
- [2] Bea, J.A. and P.C. Marijuán. 2003. The informal patterns of laughter. *Entropy*, 5, 205-213.
- [3] Bickley, C. and S. Hunnicutt. 1992. Acoustic analysis of laughter. *Proc. Int. Conf. Spoken Language Process*, 2, 927-930.
- [4] Boeke, J.D. 1891. Mikroskopische Phonogrammstudien. *Pflügers Arch. f. d. ges. Physiol.*, 50, 297-318.
- [5] Boersma, P. 1993. *Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound*, in *IFA Proceedings of the Institute of Phonetic Sciences*. p. 97-110.
- [6] Boersma, P. and D. Weenink. 2003. *Praat: Doing phonetics by computer (Version 4.02.04)*, <http://www.praat.org>.
- [7] Childers, D.G. 1978. *Modern spectrum analysis*. 1978, New York: IEEE Press.
- [8] Edmonson, M.S. 1987. Notes on laughter. *Anthropological Linguistics*, 29(1), 23-34.
- [9] Ekman, P. 1997 *What we have learned by measuring facial behavior*, in *What the face reveals*, P. Ekman and E.L. Rosenberg, Editors. Oxford University Press: New York. p. 469-485.
- [10] Fry, W.F. and C. Hader. 1977. The respiratory components of mirthful laughter. *Journal of Biological Psychology*, 19, 39-50.
- [11] Hillenbrand, J., et al. 1995. Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America*, 97(5), 3099-3111.
- [12] Joliveau, E., J. Smith, and J. Wolfe. 2004. Tuning of the vocal tract resonance by sopranos. *Nature*, 427, 116.
- [13] Kienast, M. 2002. *Phonetische Veränderungen in emotionaler Sprechweise*. 2002, Aachen: Shaker Verlag.
- [14] La Pointe, L.L., D.M. Mowrer, and J.L. Case. 1990. A comparative acoustic analysis of the laugh responses of 20- and 70-year-old males. *International Journal of Aging and Human Development*, 31(1), 1-9.
- [15] Laukkanen, A., E. Björkner, and J. Sundberg. 2004. Throaty voice quality: subglottal pressure, voice source, and formant characteristics. *Journal of Voice*, 20(1), 25-37.
- [16] Milford, P.A. 1980. *Perception of laughter and its acoustical properties*, in *Department of Speech Communication*. Pennsylvania State University: Pennsylvania.
- [17] Mowrer, D.E., L.L. LaPointe, and J. Case. 1987. Analysis of five acoustic correlates of laughter. *Journal of Nonverbal Behavior*, 11(3), 191-199.
- [18] Nwokah, E.E., et al. 1993. Vocal affect in three-year-olds: a quantitative acoustic analysis of child laughter. *Journal of the Acoustical Society of America*, 94(6), 3076-3090.
- [19] Nwokah, E.E., et al. 1999. The integration of laughter and speech in vocal communication: a dynamic systems perspective. *Journal of Speech, Language, and Hearing Research*, 42, 880-894.
- [20] Peterson, G.E. and H.L. Barney. 1952. Control methods used in a study of the vowels. *The Journal of the Acoustical Society of America*, 24(2), 175-184.
- [21] Press, W.H., et al. 1992. *Numerical Recipes in C: the art of scientific computing*. 2nd ed. 1992, New York: Cambridge University Press.
- [22] Provine, R.R. and L.Y. Young. 1991. Laughter: a stereotyped human vocalization. *Ethology*, 89, 115-124.
- [23] Rothgänger, H., et al. 1998. Analysis of laughter and speech sounds in Italian and German students. *Naturwissenschaften*, 85, 394-402.
- [24] Ruch, W. 1993 *Exhilaration and humor*, in *Handbook of emotions*, M. Lewis and J.M. Haviland, Editors. Guilford Press: New York. p. 605-616.
- [25] Ruch, W. and P. Ekman. 2001 *The expressive pattern of laughter*, in *Emotion, qualia, and consciousness*, A. Kaszniak, Editor. Word Scientific Publisher: Tokyo. p. 426-443.
- [26] Sundberg, J. and J. Skoog. 1997. Dependence of jaw opening on pitch and vowel in singers. *Journal of Voice*, 11(3), 301-306.
- [27] Szameitat, D.P. 2006. *Perzeption und akustische Eigenschaften von Emotionen in menschlichem Lachen*. *Unpublished PhD thesis*. University of Tübingen: Tübingen.