Date deposited: 7th May 2010

Version of file: Author final

Peer Review Status: Peer Reviewed

Citation for published item:

Szameitat DP, Alter K, Szameitat AJ, Wildgruber D, Sterr A, Darwin CJ. Acoustic profiles of distinct emotional expressions in laughter. *Journal of the Acoustical Society of America* 2009,**126** 1 354-366.

Acoustic profiles of distinct emotional expressions in laughter

Diana P. Szameitat[a]
Department of Psychiatry and Psychotherapy, University of Tübingen, Osianderstrasse 24, 72076 Tübingen, Germany and Department of Psychology, University of Sussex, Brighton BN1 9QG, United Kingdom

Kai Alter
Institute of Neuroscience, Newcastle University, Framlington Place, Newcastle NE1 4HH, United Kingdom

André J. Szameitat
Department of Psychology, Ludwig Maximilians University, Leopoldstrasse 13, 80802 Munich, Germany

Dirk Wildgruber
Department of Psychiatry and Psychotherapy, University of Tübingen, Osianderstrasse 24, 72076 Tübingen, Germany

Annette Sterr
Department of Psychology, University of Surrey, Guildford GU2 7XH United Kingdom

Chris J. Darwin
Department of Psychology, University of Sussex, Brighton BN1 9QG United Kingdom

(Received 16 July 2008; revised 22 April 2009; accepted 24 April 2009)

Although listeners are able to decode the underlying emotions embedded in acoustical laughter sounds, little is known about the acoustical cues that differentiate between the emotions. This study investigated the acoustical correlates of laughter expressing four different emotions: joy, tickling, taunting, and schadenfreude. Analysis of 43 acoustic parameters showed that the four emotions could be accurately discriminated on the basis of a small parameter set. Vowel quality contributed only minimally to emotional differentiation whereas prosodic parameters were more effective. Emotions are expressed by similar prosodic parameters in both laughter and speech.

[DOI: 10.1121/1.3139899]

PACS number(s): 43.71.Bp, 43.70.Gr, 43.72.Ar [DOS] Pages: 1–XXXX

I. INTRODUCTION

Laughter is a prominent part of human non-verbal communication; in social interaction it is uttered in a wide variety of different situations and emotional contexts.[1,2] Moreover, while its acoustical signal is easily identifiable,[3] it is also extremely variable.[4] Such variability is not random but, amongst other things, allows listeners reliably to perceive which of a number of different emotions is being expressed.[5] However, we do not know what acoustic properties of laughter cue the different emotions. The aims of the current study are to describe the acoustical properties of laughter sounds produced under different emotions and to test for differences between them.[6]

To our knowledge, previous studies on the acoustical structure of laughter investigated laughter emitted in single behavioral contexts.[4,8,9] However, studies directly comparing different laughter types are lacking. Thus, we derived hypotheses for acoustic cues conveying emotions in laughter from studies on emotions in speech. Numerous studies have shown that emotions are not predominantly communicated via lexical information but rather via emotional prosody (for reviews see Refs. 10–12). Different emotions in speech can be reliably identified via a small set of prosodic vocal parameters[11] such as fundamental frequency (F0), standard deviation of F0, intensity, duration of voiced elements, and energy below 1000 Hz.[12] These parameters are not unique to speech: emotional expression in musical performance is based on the same vocal indicators as has been reported for emotional speech prosody.[10] In addition, there is some evidence that similar effects are seen in non-verbal utterances[13,14] such as crying or screaming and in interjections (e.g., “yippee!” and “hurray!”). Thus, communication of emotions may rely on similar acoustic parameters in these different types of utterance.

In order to investigate emotional expressions in laughter, we analyzed four different portrayals of laughter sounds. First, we decided to test joyous and taunting laughter, as both arise from basic emotions[15] which have been regularly investigated in emotional facial and vocal expression and which differ strongly from each other.[5,13] Joyful laughter is based on joy, which resembles a positive emotion for both sender and listener, and promotes social bonding. In contrast, taunting laughter (which we consider to be synonymous to sneering laughter) is based on an aggressive, destructive emotion such as contempt or scorn, which humiliates the listener and segregates members from group context.[5] The third emotion we investigated was schadenfreude (pleasure in another’s misfortune), which resembles an affect blend of taunt (Ger-

---

[a] Author to whom correspondence should be addressed. Electronic mail: d.szameitat@gmx.de

man "schaden" = English harm) and joy (German "freude"). Although schadenfreude shares features with both, joyful and taunting laughter, it can be distinguished from the latter two emotions. Schadenfreude is similar to joy in that the sender enjoys the situation which is the misfortune of the other person. However, this joy does not (in contrast to joyful laughter) promote social bonding. Furthermore, and comparable with taunting laughter, schadenfreude aims at dominating the other person.[5] However, in schadenfreude (in contrast to taunt) the sender neither wants to seriously harm the listener nor to segregate him from the group structure. Thus schadenfreude shares similarities with teasing, a behavior that is also found in other social contexts such as between friends and romantic couples.[16–18] The fourth laughter type we tested was laughter provoked by tickling (hereafter named tickling laughter), which is one of the first laughter expressions in children[19] and one of the very few laughter expressions also emitted by non-human primates.[20,21] It is still a matter of debate whether tickling laughter is based on an emotion[22] or if it is merely a reflex action[23] (however, for ease of reading we will subsume it under the category of emotional laughter). Tickling laughter is characterized by a high physical activation and, like joyful laughter, promotes social relationships.[22]

In order to allow for a good acoustical differentiation, we analyzed the laughs according to the three basic perceptual dimensions of vocal sounds, i.e., frequency, tempo, and intensity.[24,25] Scherer[12] suggested that differentiation between emotions may be hampered if too few acoustical parameters are investigated. Accordingly, we investigated a broad range of parameters for each perceptual dimension. This also allowed for a better comparison of our data with previously reported acoustical data on emotional vocal expressions, as previously investigated parameter sets were heterogeneous. Furthermore, we examined parameters characterizing voice quality, such as amount of voiced energy, as they are essential for characterizing emotions in the human voice[26] and for differentiating laughs.[27] In order to investigate a possible contribution of vowel quality to the encoding of emotions in laughter, further analyses dealt with potential phonological content in laughter.

If emotions in laughter are communicated via similar parameters to those expressing emotions in speech, we would expect that joyful laughter is characterized by a high laugh rate, high F0, and high intensity, similar to joyful speech,[9,28,29] while taunting laughter is characterized by a low laugh rate, low F0, and a low intensity, similar to taunting speech.[28,30–35] For schadenfreude and tickling laughter, no hypothesis could be derived as their emotional speech prosody has not yet been investigated.

II. METHOD

A. Data collection

For the portrayals of emotional laughter eight professional actors (three male) produced four types of laughter, i.e., joyous, tickling, schadenfreude, and taunting. The speakers were instructed to put themselves into the respective emotional state with the help of self-induction techniques and to laugh freely without thinking about the expression of the laughter. Instructions included an example scenario for each emotion; however, the interpretation and expression of the emotions was left to the speakers to decide for themselves (see Ref. 36 for a similar approach).

TABLE I. Number of laughter sequences per speaker and emotion. ma-mc *male speakers*, fa-fe *female speakers*, J *Joy*, Ti *Tickling*, S *Schadenfreude*, Ta *Taunt*.

Speaker	J	Ti	S	Ta	Total
ma	6	1	3	1	11
mb	4	5	1	6	16
mc	6	5	6	6	23
fa	5	3	⋯	2	10
fb	4	6	2	6	18
fc	⋯	6	3	5	14
fd	5	4	4	6	19
fe	6	2	2	6	16
Total	36	32	21	38	127

Sound recordings, using a DAT recorder (TASCAM DA-P) with the microphone (Sanyo MP-101) approximately 0.5 m in front of the talker, took place in a sound proof booth. Recordings were digitized at a sampling rate of 48 kHz (16 bits), normalized, and cut into individual laughter sequences.

B. Stimulus-material

Sequences containing verbal material, interjections, and background noise were excluded from further analysis. Furthermore, only the laughter sequences that gave good expression of the emotions in a previous study[5] were used. This study divided 429 sequences into three subsets (120–153 sequences each). Each subset was then classified according to the underlying emotion in a four-choice classification paradigm by 24 (12 male) English native subjects (mean age 22 years, total $n = 72$).[5] From all correctly classified sequences (i.e., classification above chance level, $p < 0.05$, two-tailed), a stimulus set was chosen which was balanced with respect to emotion, speaker sex, and speaker identity. This set consisted of 127 laughter sequences (21–38 per emotion, 0–6 per emotion and speaker, Table I) and had an average correct classification rate of 63% (for details see Table II).

TABLE II. Classification results in percent as derived by listener's classification (Ref. 5). J *Joy*, Ti *Tickling*, S *Schadenfreude*, Ta *Taunt*. Bold type represents correct classification.

		Response			
		J	Ti	S	Ta
Stimulus	J	**61**	12	21	5
	Ti	13	**68**	15	4
	S	22	11	**54**	14
	Ta	6	4	20	**70**

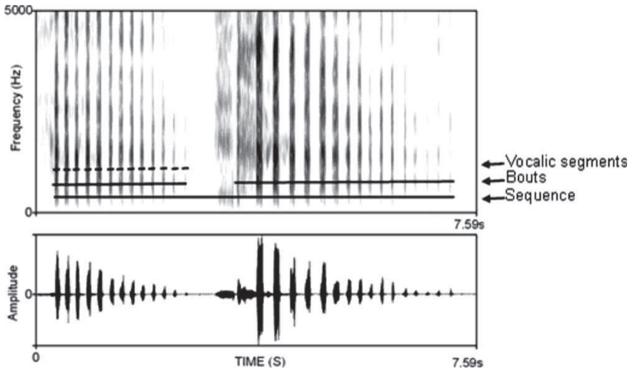

FIG. 1. Segmentation of a laughter sequence. Shown are the spectrogram (above) and oscillogram (below).

C. Acoustical analysis

The acoustic parameters were extracted using PRAAT 4.02.04.[37] Laughter sequences were segmented in the time domain according to vocalic segments (burst of energy of unvoiced and voiced exhaled breath having a single vocal peak) and bouts (either all segments from the first to the beginning of an inhaled breath or all segments between two inhaled breaths, Fig. 1). The boundaries of a segment were determined visually in the amplitude-time spectrum (distinct rise of energy from background noise into a single vocal peak) and transcribed into a script (Text-Grid function in PRAAT). On the basis of this segmentation, 43 acoustical parameters were calculated by PRAAT scripts for each individual sequence (Table III). To calculate the amplitude parameters, the values of the sounds were squared and convolved with a Gaussian window (Kaiser-20, side lobes below −190 dB, e.g., intensity: get mean function). Parameters of fundamental frequency were determined by an autocorrelation method [e.g., sound: to pitch (ac) function]. To avoid artifacts in F0 extraction, the F0 search range (pitch floor and pitch ceiling) was determined by visual inspection, i.e., by overlaying the automatically extracted pitch contours with a narrowband fast Fourier transform (FFT)-based spectrogram (30 ms, Gaussian window, pre-emphasis +6 dB/octave). For male speakers the F0 search range was always 75–600 Hz. For female speakers the F0 search range was highly variable; although it predominantly had an average range of 120–1000 Hz, the pitch ceiling could be as high as 2000 Hz. Formants were extracted by linear predictive coding [Gaussian-like window, Formant (burg) function],[38,39] a short-term spectral analysis approximating the spectrum of each analysis frame by five formants. The ceiling of the formant search range for the first five formants was 5000 Hz for male speakers and 5500 Hz for female speakers, respectively. For vocalic segments 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). The harmonic-to-noise ratio (HNR) was calculated by a short-term HNR analysis performing an acoustic periodicity detection on the basis of a forward cross-correlation analysis [harmonicity (cc) function] with a time resolution of 10 ms. The parameters center of gravity (COG), kurtosis, and skewness were calculated on the basis of the averaged spectrum [spectrum (fft) function].

AQ: #1

For calculation of parameters based on vocalic segments (segment parameters, see Table III) acoustical measurements from laughter segments that were produced with a closed mouth, or where spectral measurement extraction was uncertain were excluded leaving 3947 (125) of the original 4238 (127) laughter segments (sequences) for analysis.

D. Statistical analysis

1. Parameter-wise analysis

To test if individual acoustical parameters differed between the emotions, individual analyses of variance (ANOVAs) were calculated for each of the 43 acoustical parameters.

In detail, for parameters based on laughter sequences (sequence parameters, see Table III) some parameters were averaged across bouts (averaged: N_Sg_Bt, BtDur, IntBtDur; not averaged: TotDur, N_Sg, N_Bt, LgRate). Next, individual two-factorial ANOVAs (*emotion* (4), *speaker sex* (2), Bonferroni-corrected for 43 comparisons: overall $p < 0.05$, i.e., individual alpha level=0.0012) were carried out. Additionally, pairwise comparisons between all four emotions were calculated for each acoustical parameter showing a significant effect of emotion using Tukey's HSD tests (corrected for six comparisons).

For the evaluation of the segment parameters (see Table III) careful consideration of the acoustical properties of the laughter signal is necessary in order to avoid artifacts in the statistical analysis. For instance, the average number of vocalic segments in the sequence differed significantly between emotions [one factorial ANOVA, $F(3,117)=3.731$; $p < 0.05$]. In addition, for 20 of the segment parameters the factor *segment position* was significant (one factorial ANOVA, all $p < 0.05$, not corrected for multiple comparisons), indicating that many parameters change along the course of the laughter sequence. These two effects together might lead to artifacts in the statistical analysis. For example, two types of laughter may show a statistically significant difference with respect to the mean (averaged across segments) of a parameter that has a gradient of continually decreasing values along the laughter sequence (such as F0), although the true gradients of both laughter types are identical and the laughter types differ solely in the number of segments per bout.

In the same way, testing whether parameters change along the segments of bouts is complicated by the fact that the first segment was significantly longer than all following segments (mean duration first segment=129 ms, second segment=102 ms, Tukey-HSD contrasts for one factorial ANOVA, factor *segment position* (6), segments 1 vs 2, $p < 0.001$; for all other combinations of segments 2–6 nanoseconds) and 32 segment parameters correlated significantly with segment duration (Pearson's correlation coefficient, two-tailed, $n=1058–3932$, all $p < 0.05$). Changes in a parameter with segment number may arise simply because the first segment is longer, and the parameter

TABLE III. Investigated acoustical parameters. Parameters marked with (+) were subjected to the discriminant analysis.

Parameter	Abbreviation	Unit	Description
Sequence level			
Number of vocalic segments	N_Sg		Number of segments
Number of bouts	N_Bt		Number of bouts (separated by inbreath)
Segments per bout	N_Sg_Bt		Average number of segments in bout
Total duration	TotDur	ms	Duration from onset to end of sequence
Bout duration$^{(+)}$	BtDur	ms	Average duration of laughter bouts
Inter bout duration	IntBtDur	ms	Average duration between bouts
Laugh rate$^{(+)}$	LgRate	1/s	Average number of segments per second
Segment level			
Duration			
Segment duration$^{(+)}$	SgDur	ms	Average duration of a segment
Inter segment duration	IntSgDur	ms	Average duration between the end of a segment to the start of the following segment within a bout
Event duration	EvntDur	ms	Average duration between the start of two consecutive segments within a bout, (SgDur+IntSgDur)
Amplitude (Amp)			
Amplitude ratio	AmpMN_Max		Ratio of mean intensity to maximal intensity; (mean Amp./maximal Amp.)
Amplitude bandwidth$^{(+)}$	AmpBW	dB	Difference between maximal intensity and minimal intensity, (maximal Amp. − minimal Amp.)
Amplitude SD ratio	AmpSD_MN		Ratio of intensity standard deviation to mean intensity, (Amp. SD/mean Amp.)
Time of max. amplitude	tiAmpMax	ms	Relative position of max. Amp. measured from voice onset of segment
Fundamental frequency (F0)			
Mean F0$^{(+)}$	F0MN	Hz	Average fundamental frequency measured across time segments (i).
Minimal F0	F0Min	Hz	$\mathrm{F0Min}=\mathrm{Minimum}\ (\mathrm{F0}_i : 1 \le i \le N)$
Maximal F0	F0Max	Hz	$\mathrm{F0Max}=\mathrm{Maximum}\ (\mathrm{F0}_i : 1 \le i \le N)$
F0 bandwidth	F0 BW	Hz	$\mathrm{F0BW}=\mathrm{F0Max}-\mathrm{F0Min}$
F0start	F0Start	Hz	$\mathrm{F0}_{i=1}$
F0 end	F0End	Hz	$\mathrm{F0}_{i=N}$
F0 change	F0Chg	Hz	$\mathrm{F0Chg}=\mathrm{F0End}-\mathrm{F0Start}$
Time of max F0	tiF0Max	ms	Relative position of max. F0 measured from voice onset of segment
Formants			
F1$^{(+)}$, F2$^{(+)}$, F3, F4, F5	F1–F5	Hz	First to fifth formant
F1 bandwidth	BwF1	Hz	Bandwidth of first formant
Peak frequency (PF)			
Mean PF	PFMN	Hz	Average peak frequency measured across time segments (i).
Maximal PF$^{(+)}$	PFMax	Hz	$\mathrm{PFMax}=\mathrm{Maximum}\ (\mathrm{PF}_i : 1 \le i \le N)$
Ratio mean PF/mean F0	PFMN_F0MN		Ratio mean PF to mean F0
Ratio max PF/mean F0$^{(+)}$	PFMax_F0MN		Ratio maximal PF to mean F0
Time of max. PF	tiPFMax	ms	Relative position of max. PF measured from voice onset of segment
Voice parameters			
Ratio of voiced elements$^{(+)}$	% voic	%	Percent of time segments which had a clear harmonic structure
Mean harmonic-to-noise ratio (HNR)$^{(+)}$	HNRMN		Average HNR
HNR SD	HNRSD		Standard deviation of HNR
Maximal HNR	HNRMax		Peak HNR
Time of max HNR	tiHNRMax	ms	Relative position of max. HNR measured from voice onset of segment
Jitter	Jitt	%	Measure for micro irregularities in F0

TABLE III. (*Continued.*)

Parameter	Abbreviation	Unit	Description
Sequence level			
Shimmer	Shim	%	Measure for micro irregularities in amplitude of F0
Center of gravity$^{(+)}$	CoG	Hz	Frequency at which the energy of the signal is divided into half. Measure for the average height of the frequencies in the segment.
Skewness	Skew		Normalized skewness is the third central moment divided by the 1.5 power of the second central moment. Measure for how much the shape of the spectrum below the CoG is different from the shape above the CoG.
Kurtosis	Kurt		Normalized kurtosis is the fourth central moment divided by the square of the second central moment. Measure for how much the shape of the spectrum around the CoG is different from a Gaussian curve.

changes with segment duration rather than with segment number. This problem also prevents us from saying whether such changes differ across emotions.

Different segment positions also had different sample sizes, whereby the sample size decreased with increasing segment position, with the exception of the first segment which had a smaller sample size than the second segment. A smaller sample size, however, might result in a less accurate estimate of the mean. For the examination of the segment parameters only segments with a sample size of at least 50% of the second segment were examined, which was true for all segments up to the eighth segment. Furthermore, due to the above mentioned particularities, the first segment was excluded from the analysis.

To test whether the average value of segment parameters differed between the emotions, the parameter values for segments 2–8 were first each averaged across bouts. These seven averaged values were then themselves averaged across segments resulting in one data point per sequence for each acoustical parameter. Individual two-factorial ANOVAs were carried out on these values [*emotion* (4) × *speaker sex* (2), Bonferroni-corrected for 43 comparisons] for each parameter. Furthermore, for each parameter pairwise comparisons of the emotions were conducted using Tukey's HSD test (corrected for six comparisons).

2. Variation of parameters along bouts

To test for parameter changes during the segments of a bout, the values for each of segments 2–8 were separately averaged across bouts, so that for each laughter sequence there was one data point for each of segments 2–8. Individual three-factorial ANOVAs [*emotion* (4) × *speaker sex* (2) × *segment position* (7)] were then carried out and the factor *segment position* was examined for significance (Bonferroni-corrected for 36 comparisons). To test if emotions differ in the change of parameters along the bouts, we examined, in a second step, the interaction *segment position* × *emotion* (Bonferroni-corrected for 36 comparisons). To understand potential interactions more thoroughly, we calculated, separately for each parameter, all pairwise combinations of emotions in separate ANOVA [*emotion* (2) × *segment position* (7)]. Finally, to test for the direction of potential parameter changes along the bouts, we calculated a linear regression for each parameter and emotion.

3. Analysis of the first segment

The above statistical analysis used only the second to eighth segments. To test whether the first segment contains further information for differentiating between emotions beyond the one provided by segments 2–8 further analysis was made to test differences between the first and second segments. Parameter values for segments 1 and 2 were separately averaged across bouts and individual three factorial ANOVA performed [*emotion* (4) × *speaker sex* (2) × *segment position* (2), Bonferroni-corrected for 36 comparisons]. A significant interaction between the factors *emotion* and *segment position* would indicate that differentiation of emotions depends on the segment. Further analysis will be conducted for such parameters to test whether the first segment provides information beyond the one carried by the second segment.

4. Identification of emotions

To test how well different emotions can be identified, a subset of acoustical parameters was subjected to a discriminant analysis (Table III). Parameters were chosen according to the following criteria: First, at least one parameter was chosen from each parameter domain [domains: (1) sequence parameter in general, on the segment level: (2) duration, (3) amplitude, (4) fundamental frequency, (5) formants, (6) peak frequency, (7) voice parameters, see Table III]. Second, only parameters showing significant differences between the emotions (individual two-factorial [*emotion* (4) × *speaker sex* (2)] ANOVA, $p < 0.05$, Bonferroni-corrected for 43 comparisons) were selected, with the exception of the parameter bout duration, which was included since it missed the significance level only by a small margin ($p = 0.0013$ instead of the required $p < 0.0012$ for $p < 0.05$, Bonferroni-corrected for 43 comparisons). Finally, we predominantly chose parameters which did not correlate with any other parameter. However, following Hammerschmidt and Jürgens,[26] we retained some correlated parameters which both theoretical considerations and empirical findings deemed important for characterizing prosodic structure. To assess the discriminative power of each individual parameter, we additionally calculated 12 separate discriminant analysis, one for each parameter.

5. Vowel quality

To identify the vowel quality of vocalic segments, F1-F2 plots were generated and compared with the standard vowel space representation according to Hillenbrand *et al.*[40] To examine if emotions are characterized by specific vowels, F1-F2 plots were compared with emotion recognition rates for each talker.

III. RESULTS

A. Differentiation of individual parameters

To examine the acoustical correlates of laughter sounds expressing different emotions, we first tested whether individual acoustical parameters differed between the emotions by conducting 43 individual two-factorial ANOVA [*emotion* (4) × *speaker sex* (2)]. This analysis revealed that 26 out of 43 investigated parameters differed significantly between the four emotions (all $p<0.05$, Bonferroni-corrected, $F(42)=5.885-50.734$, Table IV). For sequences, the parameters: number of bouts (N_Bt), temporal distance between bouts (IntBtDur), and laugh rate (LgRate) differed. For segments, two duration parameters (SgDur, EvntDur), many amplitude parameters (AmpBW, AmpSD_MN, tiAmpMax), most F0 parameters (F0MN, F0Min, F0Max, F0BW, F0Start, F0End), the first and second formants (F1, F2), all peak frequency parameters (PFMW, PFMax, PFMW_F0, PFMax_F0, tiPFMax), % of voiced elements, mean HNR, CoG, skewness, and kurtosis differed significantly between the emotions. Thus, the different laughter types clearly had different acoustical properties.

Additional analyses revealed that 21 acoustical parameters showed differences between male and female speakers (factor *speaker sex*, all $p<0.05$). The laughter of female speakers had higher frequencies (F1-F5, CoG, all F0 and PF parameters with the exception of F0Chg, tiPFMax) was more regular and more voiced (jitter, shimmer, HNR, % voiced elements), and the time of F0max measured from voice onset was longer (tiF0max). Moreover, six of the acoustical parameters showing differences between the emotions had a significant interaction between the factors *emotion* and *speaker sex* (EvntDur, F0MN, F0Min, F0Max, F0BW, F0Start, all $p<0.05$): male and female speakers thus modulated some parameters differently.

B. Differentiation of changing patterns of individual parameters

There was significant change along the course of the bout for 15 of the 36 segment parameters (three factorial ANOVA [*emotion* (4) × *speaker sex* (2) × *segment position* (7)], factor *segment position*, all $p<0.05$, Bonferroni-corrected). The segment duration, many F0 parameters (F0MN, F0Min, F0Max, F0BW), some voice parameters (%voic, HNRMW, HNRSD), and one amplitude parameter (AmpMN_Max) decreased along bouts, while the ratio between PF and F0 (PFMW_F0, PFMax_F0), jitter and shimmer, and two amplitude parameters (AmpBW, AmpSD_MN) increased along bouts. However, only one parameter (PFMax_F0) showed a different pattern of change depending on the emotion (interaction *segment position* × *emotion*, $p<0.05$). This interaction was due to PFMax_F0 increasing more with increasing segment position in taunt than in joy or tickling laughter [individual three-factorial ANOVAs (*emotion* (2) × *speaker sex* (2) × *segment position* (7)], interaction *emotion* (taunt vs joy or taunt vs tickling, respectively) × *segment position*, $p<0.05$; linear regressions (all $p<0.05$): PFMax_F0: β taunt=0.32, β joy=0.10, β tickling =0.22). These results indicate that the pattern of parameter changes along the bout contributes only minimally to the differentiation of emotions.

C. The first segment

To test whether the first segment provides further information for acoustical differentiation beyond the one derived from the analysis of segments 2–8, we tested in individual three-factorial ANOVAs [*emotion* (4) × *speaker sex* (2) × *segment position* (2)] if the first and second segments (averaged across bouts) differed acoustically. A significant interaction between the factors segment and emotion was evident only for two acoustical parameters (both $p<0.05$, Bonferroni-corrected), i.e., % of voiced elements (%voic) and CoG. In detail, in joyous laughter the percentage of voiced elements was lower in the first than in the second segment, while there were no differences between the first and second segments for tickling, taunt, and schadenfreude. The CoG showed the opposite pattern for joy, since the first segment had higher values than the 2nd segment, while the 1st and 2nd segment did not differ for tickling, taunt, and schadenfreude. However, visual inspection of this pattern indicated that the differences between the emotions were larger in the second segment as compared to the first segment. Therefore, we suggest that the first segment adds only little additional information for the differentiation of emotions expressed in laughter.

D. Identification of emotions

To test how well different emotions can be identified, a discriminant analysis was conducted on the basis of a reduced parameter set. Acoustical parameters where chosen according to the following criteria: parameters which (1) described different acoustical cues, (2) differed significantly and strongly (high *p*-value) between the emotions, and (3) showed little correlation (for details see Sec. II D 4). The resulting parameter set consisted of the following 12 acoustical parameters: F0, F1, F2, SgDur, MaxPF_F0, MaxPF, AmpBW, %voic, HNRMN, CoG, BtDur, and LgRate (Table III). We found that the emotional category of the laughter stimuli could be predicted with a high accuracy (discriminant analysis "enter-method" ("leave-one out cross validation"): mean 84% (76%), for details see Table V).

TABLE IV. Mean values for the four types of laughter and results of statistical tests. Pairwise t-tests were calculated for all combinations of laughter type [e.g., J-Ti *pairwise t-test joy vs tickling*, left arrows (〈) joy significantly smaller than tickling, right arrows (〉) joy significantly higher than tickling; all other comparisons equivalent]. (〈, 〉) $p < 0.05$, (〈〈, 〉〉) $p < 0.01$, (〈〈〈, 〉〉〉) $p < 0.001$. Abbreviations. Sex *speaker sex*, F *female speakers*, M *male speakers*, J *Joy*, Ti *Tickling*, S *Schadenfreude*, Te *Taunt*. For further abbreviations and units of acoustical parameters see Table III.

Parameter sequence level	Sex	Means					t-tests					
		J	Ti	S	Ta	Total	J-Ti	J-S	J-Ta	Ti-S	Ta-Ti	Ta-S
NrSg	F	32.5	30.7	33.9	30.3	31.5						
	M	31.7	42.2	33.9	38.8	36.2						
NrBt	F	3.0	4.3	3.5	3.3	3.5	〈〈〈			〉〉	〈〈	
	M	2.8	4.6	2.8	3.4	3.4						
NrSg_Bt	F	13.1	7.6	10.5	9.5	10.1						
	M	12.5	11.1	12.3	11.3	11.8						
TotDur	F	7940	6749	7685	7376	7404						
	M	7540	8826	9029	8778	8436						
BtDur	F	2644	1390	2034	1945	1996						
	M	2481	1976	3018	2291	2431						
IntBtDur	F	698	329	439	515	498	〉〉〉	〉〉〉	〉〉〉	〈〈	〉	
	M	783	419	628	474	590						
LgRate	F	4.08	4.60	4.38	4.07	4.26	〈〈〈			〉〉	〈〈	
	M	4.20	4.87	3.77	4.33	4.29						
Segment level												
Duration												
SgDur	F	88	82	90	109	94			〈〈	〈〈	〉〉〉	
	M	90	85	116	101	97						
IntSgDur	F	114	105	112	107	109						
	M	123	100	144	113	120						
EvntDur	F	202	189	204	217	204					〉〉〉	
	M	214	185	259	214	217		〈		〈〈〈		〈
Intestity												
AmpMN_Max	F	0.928	0.913	0.912	0.898	0.912						
	M	0.918	0.922	0.907	0.914	0.916						
AmpBW	F	0.250	0.305	0.299	0.369	0.311		〈	〈〈〈		〉〉	
	M	0.266	0.251	0.310	0.291	0.278						
AmpSD_MN	F	0.081	0.099	0.100	0.120	0.101		〈	〈〈〈		〉〉	
	M	0.093	0.090	0.106	0.102	0.097						
tiAmpMax	F	44	42	49	60	49			〈〈	〈	〉〉〉	
	M	48	48	61	53	52						
Fundamental frequency												
F0MN	F	500	681	412	329	479	〈〈〈		〉〉〉	〉〉〉	〈〈〈	
	M	177	261	216	158	199		〈		〉		〈〈
F0Min	F	431	599	366	296	421	〈〈〈		〉〉〉	〉〉〉	〈〈〈	
	M	154	237	189	148	178				〉		〈
F0Max	F	547	744	445	354	521	〈〈〈		〉〉〉	〉〉〉	〈〈〈	
	M	198	279	243	164	217		〈				〈〈〈
F0BW	F	117	146	79	58	100			〉〉〉	〉〉〉	〈〈〈	
	M	44	41	55	16	39					〈〈	〈〈〈
F0Start	F	481	713	430	331	485	〈〈〈		〉〉〉	〉〉〉	〈〈〈	
	M	198	268	252	157	215		〈〈	〉			〈〈〈
F0End	F	447	604	394	294	432	〈〈〈			〉〉〉	〈〈〈	
	M	144	252	178	116	177						
F0Chg	F	36	51	21	7	30						
	M	49	30	65	29	44						
tiF0Max	F	51	42	48	53	49						
	M	25	30	42	34	32						
Jitt	F	0.03	0.02	0.03	0.02	0.02						
	M	0.05	0.04	0.03	0.03	0.04						
Shim	F	0.12	0.13	0.15	0.14	0.13						
	M	0.24	0.21	0.19	0.18	0.21						
Formants												

TABLE IV. (*Continued.*)

Parameter sequence level	Sex	Means					t-tests					
		J	Ti	S	Ta	Total	J-Ti	J-S	J-Ta	Ti-S	Ta-Ti	Ta-S
F1	F	802	909	967	1052	936		⟨⟨	⟨⟨⟨		⟩⟩⟩	
	M	660	654	797	829	728						
F2	F	1654	1736	1666	1745	1707	⟨⟨⟨		⟨	⟩⟩		
	M	1462	1686	1485	1500	1526						
F3	F	2962	2907	3011	3027	2976						
	M	2666	2767	2685	2649	2688						
F4	F	3800	3757	3878	3913	3837						
	M	3523	3449	3603	3314	3471						
F5	F	4578	4661	4629	4604	4616						
	M	4205	4262	4240	4147	4211						
BwF1	F	153	172	241	155	171						
	M	192	157	164	122	161						
Peak frequency												
PFMW	F	870	1049	1077	1179	1049	⟨	⟨	⟨⟨⟨			
	M	540	672	822	890	713						
PFMax	F	856	1018	1195	1285	1089		⟨⟨⟨	⟨⟨⟨		⟩⟩⟩	
	M	649	715	917	943	791						
PFMW_F0	F	1.8	1.6	2.9	3.8	2.6		⟨⟨	⟨⟨⟨	⟨⟨⟨	⟩⟩⟩	⟩
	M	3.0	2.3	4.1	6.1	3.9						
PFMax_F0	F	1.7	1.5	3.2	4.2	2.7		⟨⟨⟨	⟨⟨⟨	⟨⟨⟨	⟩⟩⟩	⟩⟩
	M	3.5	2.5	4.5	6.5	4.2						
tiPFMax	F	43	42	49	61	50			⟨⟨⟨		⟩⟩⟩	
	M	50	50	61	56	54						
Voice parameters												
%voic	F	87	82	74	66	77		⟩⟩	⟩⟩⟩	⟩	⟨⟨⟨	
	M	69	67	52	39	58						
HNRMW	F	11.2	11.4	8.7	8.3	9.9			⟩	⟩⟩	⟨⟨⟨	
	M	6.5	7.9	5.7	5.2	6.3						
HNRSD	F	4.8	5.0	4.5	5.2	4.9						
	M	4.1	4.1	4.3	4.7	4.3						
HNRMax	F	23.4	26.2	23.8	24.9	24.7						
	M	23.7	25.5	24.0	25.4	24.6						
tiHNRMax	F	44	40	48	52	46						
	M	46	47	60	45	49						
CoG	F	1163	1409	1440	1646	1427	⟨⟨	⟨⟨	⟨⟨⟨		⟩⟩	⟩
	M	804	1033	1139	1255	1034						
Skew	F	5.9	5.4	4.1	3.4	4.7			⟩⟩⟩		⟨⟨⟨	⟨
	M	6.0	4.9	5.7	3.2	5.0						
Kurt	F	92	79	52	30	62			⟩⟩⟩		⟨	
	M	95	53	85	33	68						

To test the discrimination power of each parameter individually, we calculated 12 separate discriminant analyses. These analyses revealed that emotions could be classified with an accuracy of 33.6%–48.0% (leave-one out cross validation) on the basis of a single parameter (Fig. 2).

E. Vowels

The vowel elements of the laughter sequences were predominantly based on central vowels characterized by middle F2 values, with vowel height varying from mid (ə) to open (a) (for details see Ref. 41).

To test whether vocalic elements contributed to emotional differentiation, first F1-F2 plots were analyzed for each speaker individually and then compared with the speak-

TABLE V. Classification results in percent as derived by discriminant analysis. J *Joy*, Ti *Tickling*, S *Schadenfreude*, Ta *Taunt*. Bold type represents correct classification.

		Predicted			
		J	Ti	S	Ta
"enter-method"	J	**89**	3	6	3
	Ti	3	**94**	0	3
	S	24	10	**52**	14
	Ta	0	0	11	**89**
"leave-one out cross validation"	J	**81**	3	14	3
	Ti	6	**81**	3	10
	S	29	10	**43**	19
	Ta	0	0	14	**86**

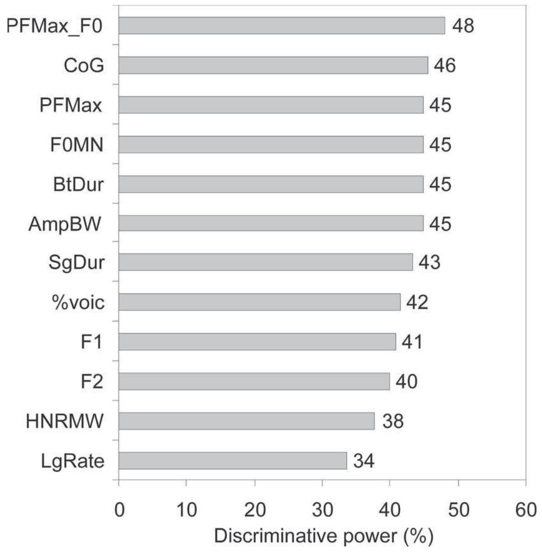

FIG. 2. Discriminative power of individual parameters. Calculated by separate discriminant analyses (leave-one out cross validation). For abbreviations of acoustical parameters, see Table III.

TABLE VI. Acoustical correlates. J *Joy*, Ti *Tickling*, S *Schadenfreude*, Ta *Taunt*; <</>> very *small/large values*; </> *small/large values*; = *middle values*; gender effect: f *females*; m *males*; bold type: *significantly different to all remaining laughter types*.

	J	Ti	S	Ta
Segment duration	=	<	>	>
Event duration	=	<	f = m **>>**	f > m =
Laugh rate	=	**>>**	=	=
Number of bouts	=	**>>**	=	=
Inter-bout duration	**>>**	**<<**	=	=
Intensity	<	<	>	>
F0	=	**>>**	=	<<
Peak frequency	<<	<	>	>>
PF/F0	<	<	=	**>>**
F1	<<	<	>	>>
F2	<	>	<	>
% voiced elements	>	>	<	<
HNR	>>	>	<	<<
Center of gravity	**<<**	=	=	**>>**
Skewness	=	=	=	**<<**
Kurtosis	=	=	=	<<

er's individual recognition rates. F1-F2 plots for individual speakers revealed that the clusters of the vowel elements overlapped widely for most of the speakers and emotions. Furthermore, the variability in vocalic elements varied strongly with speaker identity, i.e., in four speakers the vowel elements differed between the emotions, and in three speakers the vowel elements showed virtually no difference. All speakers uttered almost exclusively central vowels (e.g., ɑ or ə), and in the rare cases where non-central vowels were expressed, recognition rates remained unchanged, which indicates that vowels were not used by the listeners to differentiate between emotions.

IV. DISCUSSION

Analysis of the expression of four different emotions in laughter revealed that they differ in a variety of acoustical parameters, and that they can be classified accurately (84%) on the basis of a small parameter set. Overall, prosodic parameters provided a good basis for classification, whereas vowel quality did not differ reliably between the emotions.

A. Prosodic characteristics of the four laughter types

Laughter sequences from the four emotions used here were associated with specific acoustical correlates (Table VI). Tickling laughter was rapid and high-pitched. Its F0 reached up to 1112 Hz for females (glottal whistles up to 1765 Hz) and up to 528 Hz for males and it had the shortest segment duration, inter-bout duration, and event duration, as well as the highest laugh rate and number of bouts. Furthermore, tickling laughter had more harmonic energy (HNR, %voic) than did schadenfreude and taunting laughter. The first formant and the peak frequency were rather low, leading in combination with the high F0 to low PF_F0 values. The second formant, on the other hand, was higher than in joyful and schadenfreude laughter, and comparable to taunting laughter. The intensity parameters were rather low.

Joyful laughter was rich in low-frequency energy and had the longest time between bouts. More specifically, it had the lowest peak frequency and first formant frequency, and its energy was the most concentrated in the lower frequency range (lowest CoG). In the time domain it stood out by having the longest temporal distance between bouts (IntBtDur). Its fundamental frequency was in the middle range, which in combination with the low peak frequency, resulted in low PF_F0 values, were comparable to those of tickling laughter. Besides which, joyful laughter had a lot of harmonic energy (HNR, %voic), similar to tickling laughter. The second formant was rather low, i.e., lower than in tickling and taunting laughter. Also the intensity parameters were rather low, i.e., they were lower than in schadenfreude and taunting laughter.

Schadenfreude laughter did not show any outstanding characteristics, i.e., most of its parameters were in the middle range. Specifically, schadenfreude laughter shared features with both joyful and taunting laughter (see Table V). In the time domain schadenfreude was comparable to joyful and taunting laughter. In the intensity domain, it was comparable to taunting laughter. Moreover, while the fundamental frequency and second formant were comparable to joyful laughter, the first formant and peak frequency were comparable to taunting laughter. This resulted in that the parameter PF_F0 was in the middle range, i.e., it was higher than in joyous and tickling laughter, but lower than in taunting laughter. Additionally, schadenfreude laughter had little harmonic energy (HNR, %voic), comparable to taunting laughter.

Taunting laughter had the lowest fundamental frequency, but the highest first formant and peak frequency giving the highest PF_F0 ratio. It also had the most energy concentrated in the higher frequency range (highest CoG) but the frequency distribution parameters skewness and kurtosis were lower in comparison to the remaining three laughter types. It had a small amount of harmonic energy (HNR, %voic) and a high segment duration whereby both parameters were comparable to schadenfreude laughter. Finally, its intensity parameters were higher than in joyful and tickling laugher.

B. Emotional expressions in laughter in comparison to speech

As shown in Sec. IV A, laughter sequences from the four emotions were associated with specific acoustical correlates. The question arises whether those acoustical correlates are unique for emotional expression in laughter, or whether commonalities exist to emotional expression in speech.

A number of findings support the latter hypothesis. First, the same parameters that showed reliable differences between the laughter types have also previously been reported to distinguish different emotions in speech, including F0 and PF, HNR, amplitude bandwidth, speech rate (see laugh rate for laughter), and CoG.[26] Moreover, the acoustical correlates of joyful and taunting laughter were mainly in accordance with the theoretical predictions made for joyful and contemptuous emotional speech prosody by Scherer[11] (assuming that taunt and contempt refer to comparable emotions). Finally, the acoustic profiles for joyful and taunting laughter are very similar to the acoustic profiles of joyful and contemptuous speech prosody. (To our knowledge schadenfreude and tickling speech prosody have not been previously investigated) In detail, taunting laughter and contemptuous speech prosody were both characterized by a low mean F0 (see Refs. 26, 30–33, and 35.) and low maximal F0,[26] a low F0 bandwidth,[26,31] a long segment duration,[33,34,26] a long temporal distance of F0max measured from voice onset (tiF0Max),[26] a low amount of harmonic energy,[26] and both utterances were often produced with a "pressed" voice.[31] However, in contrast to contemptuous speech prosody, taunting laughter had an average instead of low laugh rate,[31,34] and the peak frequency was high instead of low.[26] Joyful laughter and joyful speech prosody were both characterized by a high F0 and F0 bandwidth.[10,11] Furthermore, both expressions showed decreased values for the first formant.[42] However, in contrast to joyful speech prosody, in joyful laughter the CoG was at low instead of middle[10] frequencies and the peak frequency was low instead of high.[26]

Taken together, most of the acoustical correlates for joy and taunt were in line with previous findings for the respective emotions when communicated via speech prosody. Differences in the findings may be caused by more fine-grained differences within the employed emotions.[12] Another possibility is that emotional communication in laughter and speech is not equivalent in all acoustical correlates.

C. Laughter portrayals in comparison to spontaneous laughter

Since the stimulus-material was based on laughter portrayals produced by professional actors the question arises whether such portrayals truly reflect spontaneously emitted laughs. With respect to speech literature, the majority of authors assumed such equivalence,[43,44] although some noted that emotional portrayals may overemphasize acoustical parameters so that they may be more intense and prototypical than spontaneous expressions.[45] However, a number of findings support the assumption of equivalence.

First, the majority of the acoustical parameters of our stimulus-material fell well within the range previously reported for spontaneously emitted laughs. For example, the reported fundamental frequency was in accordance with previous studies: the average F0 was 199 Hz for males [compared to a range of previously reported average F0 (Refs. 3, 4, 8, and 46–52) 126−424 Hz] and 476 for females [160−502 Hz (Refs. 3, 4, 8, 48, and 50–53)] respectively. Moreover, most of our temporal parameters were well within the range of previously reported data: mean segment duration was 95 ms in this study, (compared[3,48,49,51–53] to means of 60−370 ms), intersegment duration was 115 ms (compared[3,4,8,48,49,51,52] to means of 87−240 ms), mean bout duration was 2213 ms (compared[3,4,46,47,51–55] to means of 700−3970 ms), and mean laugh rate was 4.3 segments/s (compared[4,46–48,51,52,54] to means of 2.8–5.6). However, the mean number of segments per bout was 11 segments and therefore on the upper limit of previously reported data (compared[3,4,8,46,47,51,52,55] to means of 1.5–12.5). The relatively high number of segments per bout has probably been caused by the fact that speakers were asked to produce long laughter sequences (the stimulus-material was intended to be also used in another study requiring longer durations). Formant measurements were in accordance with previous findings,[4,50,51] with the exception of the first formant which was much higher than previously reported [this study: males (females) 728 (924) Hz; as compared to 535 (653) Hz,[4] 543 (559),[50] females 650 Hz (Ref. 50)]. Detailed analyses revealed that high F1 values were not due to an artifact in formant extraction, but most likely reflect extreme positions adopted by the vocal tract during laughter in combination with physiological constraints accompanying production of a "pressed" voice, as reported in Refs. 41. Finally, analysis of vowel quality of vocalic segments showed that most of the vowels were based on central vowels, with only occasional deviants, which is in accordance with previous findings.[4,48,51,52,56,57] Taken together, the majority of the acoustical parameters measured in this study were in accordance with previous findings.

Second, the specific acoustical correlates of the two laugh utterances joy and taunt showed many commonalities with the respective emotions in emotional speech prosody (see Sec. IV B). Finally, laugh portrayals and spontaneous laughs are very hard to tell apart, as assessed by listeners discrimination[58] as well as the laughter's acoustical structure.[59] However, to answer the question conclusively as to whether portrayals truly reflect spontaneously emitted laughter, an investigation of emotional expression in spontaneous laughter is needed.

D. Differentiations on the basis of vowel quality

Emotional laughter is sometimes, for example, in comic strips, illustrated with certain vowels, e.g., joyous laughter is depicted as /hahaha/, taunt as /hohoho/, tickling as /hihihi/, or schadenfreude as /hɛhɛhɛ/, which may indicate a contribution of vowel quality to the encoding of emotions in laughter. However, vowel quality contributed only minimally to the discrimination of emotions in laughter, since laughter sequences were almost exclusively based on central vowels and the rare use of non-central vowels had no significant influence on the recognition rate.

Another hypothesis relating vowel quality with emotion was suggested by Ruch and Ekman.[23] They suggested that during the production of "reflexlike" laughter the vocal tract remains in a neutral position so that such laughs are not articulated, while emotional laughter would involve supralaryngeal structures leading to a diversity in vowel elements. However, our data did not support this assumption, since tickling laughter, which could be interpreted as a reflexlike laughter type, showed the same vowel elements as schadenfreude and taunt, i.e., (ə), (a), and (ɑ) vowels. In contrast, joyful laughter tended to involve more (ə) vowels, which are characterized by a neutral vocal tract, than in the other laughter types. Therefore, it was not the reflexlike laughter type, i.e., tickling laughter, which was predominantly based on unarticulated vowels, but joyful laughter, an emotional laugh utterance.

E. Emotions in laughter in comparison to other non-verbal vocalizations

The question arises how laughter should be integrated in the framework of non-verbal vocalizations. Wundt[60] classified non-verbal emotional vocalizations into two categories. In the first category are primary affective vocalizations, which he described as relicts of a pre-language period, e.g., panic shrieks (German "naturlaute," primarily interjections, raw affect bursts).[59–61] In the second category are secondary affective vocalizations, which were assimilated into language, and eventually conventionalized, e.g., "yucky!" or "hooray!" (secondary interjections, affect emblems).[60–62] Scherer[62] assumed that primary affective vocalizations are direct externalizations of motor behaviors reflecting push effects, while secondary affective vocalizations are primarily influenced by socio-cultural norms reflecting pull effects.

That non-verbal vocalizations can indeed be classified into these primary and secondary vocalizations is supported by a study of Schröder.[13] In his study some non-verbal vocalizations could be classified according to the emotions solely on the basis of their transcripts (e.g., German: "igitt," "yippie"), while others could not (e.g., yawning out of boredom). Furthermore, Dietrich *et al.*[14] showed that the transition between the two categories is continuous. Therefore, non-verbal affective vocalizations can communicate emotions via the same mechanism as that known for emotional communication via speech, i.e., lexical meaning (word content) and emotional prosody. Moreover, non-verbal vocalizations can be arranged on a continuous scale, whereby primary affective vocalizations differ merely on the basis of emotional prosody, while secondary affective vocalizations can differ in both emotional prosody and lexical meaning.[14]

The question arises where laughter should be placed on this (continuous) scale. In the present study we showed that laughter is predominantly based on central vowels and therefore is foremost not articulated. Furthermore, different emotional laughs did not differ according to a systematic variation in vowel quality, which might have been served as lexical information. Moreover, laughter is estimated to be 7×10^6 years old,[63] and thus its existence predates the evolution of language.[23] Based on these findings, we suggest that laughter is a primary affective vocalization, whereby various emotional expressions differ foremost in emotional prosody.

F. Vocal expression of emotions

With regard to the origin of emotional speech prosody, an intriguing hypothesis has been suggested. With the development of human language intensive neuronal and physiological changes took place in order to enable the production and perception of speech.[64] As the production of language and non-verbal affect vocalizations is based on the same physiological structures, i.e., the vocal tract, it has been suggested that with the development of human speech neural structures subserving speech production have been superimposed upon already existing structures subserving the production of non-verbal affective vocalizations.[28] Accordingly, emotional prosody is assumed to predate language development and to derive from animal communication.[21,28] However, evidence supporting this theory is sparse, since only little is known about emotional prosody in animal communication.[65,66]

Interestingly, some marked features of laughter may provide tentative support for this theory. Laughter is inborn, evident by the fact that also deaf-blind born children laugh.[67] It emerges in babies at the age of 4 months, and thus long before language acquisition.[23,68] Also in phylogeny it predates language evolution,[63] and it is one of the few vocalizations not only uttered by humans but also by non-human primates.[21] Therefore, laughter seems to be a phylogenetically old communication signal dating back to our primate ancestors.

A comparison of emotional expression in laughter and speech reveals numerous striking commonalities. In both laughter and speech emotions are expressed by similar acoustical parameters, in particular peak frequency, F0, temporal patterns, and resonance characteristics of the vocal tract (for emotional speech prosody see Ref. 26). Even more specifically, discrete emotions, such as joy and taunt, have highly comparable acoustical correlates when expressed in laughter and in speech. In line with the idea that the same emotional prosody underlies laughter and speech, behavioral studies revealed that the classification accuracy for emotional laughter[5] falls within the range reported for emotional speech prosody.[10] Additionally, the confusion matrices derived from the classification of emotions in laughter (see Tables II and V) and speech show similar patterns, and distinct emotions are characterized by similar values in arousal, valence, and dominance in laughter and speech.[5] This striking convergence strongly supports the hypothesis that emotions are communicated via the same mechanism in laughter and speech, i.e., emotional prosody.

Thus, the existence of emotional prosody in laughter, a phylogenetically old communication signal derived from animal communication, is one of the few indications based on empirical data which support the hypothesis[28] that emotional prosody is a communication system dating back prior to the evolution of language.

V. CONCLUSIONS

The present study showed that laughter sequences from the four emotions—joy, schadenfreude, taunt, and tickling—were associated with distinct acoustical correlates. Accordingly, the present study supports the hypotheses that acoustic distinction between different types of laughter exists, and that this acoustic variability is a potent tool for communicating the sender's emotional state to the listener. Crucially, we found that acoustical correlates of emotions in laughter had much in common with emotional expression in speech, supporting a common underlying mechanism for the vocal expression of emotions. The existence of emotional expression in laughter, a non-verbal signal existing long before development of human language, provides suggestive evidence that vocal emotional expression also existed long before evolution of language. That emotional modulation in laughter is primarily based on respiration and phonation rather than on articulation (i.e., vowel quality) suggests that only little supralaryngeal modeling is involved in vocal emotional expression, and this is a finding consistent with the notion that supralaryngeal structures become only centrally involved with the production of language.

ACKNOWLEDGMENTS

This work was supported by grants from the Marie Curie Foundation and the German Research Foundation (DFG AL357/1 and WI2101/2). This work was also supported by The Lena Teague Bequest. We thank Maria Bale for her helpful comments on this paper.

[1] F. Poyatos, "The many voices of laughter—A new audible-visual paralinguistic approach," Semiotica **93**, 61–81 (1993).

[2] H. Giles and G. S. Oxford, "Towards a multidimensional theory of laughter causation and its social implications," Bull. Br. Psychol. Soc. **23**, 97–105 (1970).

[3] H. Rothgänger, G. Hauser, A. C. Cappellini, and A. Guidotti, "Analysis of laughter and speech sounds in Italian and German students," Naturwiss. **85**, 394–402 (1998).

[4] J.-A. Bachorowski, M. J. Smoski, and M. J. Owren, "The acoustic features of human laughter," J. Acoust. Soc. Am. **110**, 1581–1597 (2001).

AQ: #2 [5] D. P. Szameitat, K. Alter, A. J. Szameitat, C. J. Darwin, D. Wildgruber, S. Dietrich, and A. Sterr, "Differentiation of emotions in laughter at the behavioral level," Emotion **9**, ■ (2009).

AQ: #3 [6] When referring to the concept of emotion, we follow the definition of Scherer (Ref. 7, p. 140) who stated that an emotion is a "relatively brief episode of synchronized responses by all or most organismic subsystems to the evaluation of an external or internal event as being major significance."

[7] K. R. Scherer, "Psychological models of emotion," in *The Neuropsychology of Emotion*, edited by J. C. Borod (Oxford University Press, Oxford, 2000), pp. 137–162.

[8] J. Vettin and D. Todt, "Laughter in conversation: Features of occurrence and acoustic structure," J. Nonverb. Behav. **28**, 93–115 (2004).

[9] J. Vettin and D. Todt, "Human laughter, social play, and play vocalizations of non-human primates: An evolutionary approach," Behaviour **142**, 217–240 (2005).

[10] P. N. Juslin and P. Laukka, "Communication of emotions in vocal expression and music performance—Different channels, same code?," Psychol. Bull. **129**, 770–814 (2003).

[11] K. R. Scherer, "Vocal affect expression: A review and a model for future research," Psychol. Bull. **99**, 143–165 (1986).

[12] K. R. Scherer, "Vocal communication of emotions: A review of research paradigms," Speech Commun. **40**, 227–256 (2003).

[13] M. Schröder, "Experimental study of affect bursts," Speech Commun. **40**, 99–116 (2003).

[14] S. Dietrich, H. Ackermann, D. P. Szameitat, and K. Alter, "Psychoacoustic studies on the processing of vocal interjections: How to disentangle lexical and prosodic information?," Prog. Brain Res. **156**, 295–302 (2006).

[15] P. Ekman, "Strong evidence for universals in facial expressions: A reply to Russell's mistaken critique," Psychol. Bull. **115**, 268–287 (1994).

[16] D. Keltner, R. C. Young, E. A. Heerey, C. Oemig, and N. D. Monarch, "Teasing in hierarchical and intimate relations," J. Pers. Soc. Psychol. **75**, 1231–1247 (1998).

[17] A. Mooney, R. Creeser, and P. Blatchford, "Children's views on teasing and fighting in junior school," Educ. Res. **33**, 103–112 (1991).

[18] L. A. Baxter, "Forms and functions of intimate play in personal relationship," Human Communication Research **18**, 336–363 (1992).

[19] R. W. Washburn, "A study of the smiling and laughing of infants in the first year of life," Genet. Psychol. Monogr. **6**, 397–537 (1929).

[20] J. A. R. A. M. van Hooff, "A comparative approach to the phylogeny of laughter and smiling," in *Nonverbal Communication*, edited by R. A. Hinde (University Press, Cambridge, 1972), pp. 209–241.

[21] C. Darwin, *The Expression of the Emotions in Man and Animals*, 3rd ed., edited and republished by P. Ekman (Harper Collins, London, 1872).

[22] J. Panksepp and J. Burgdorf, "'Laughing' rats and the evolutionary antecedents of human joy?," Physiol. Behav. **79**, 533–547 (2003).

[23] W. Ruch and P. Ekman, "The expressive pattern of laughter," in *Emotion, Qualia, and Consciousness*, edited by A. Kaszniak (Word Scientific, Tokyo, 2001), pp. 426–443.

[24] J. Pittam and K. R. Scherer, "Vocal expression and communication of emotion," in *Handbook of Emotions*, edited by M. Lewis and J. M. Haviland (Guilford, New York, 1993), pp. 185–197.

[25] K. R. Scherer, "Methods of research on vocal communication: Paradigms and parameters," in *Handbook of Methods in Nonverbal Behavior Research*, edited by K. R. Scherer and P. Ekman (Cambridge University Press, Cambridge, 1982), pp. 136–198.

[26] K. Hammerschmidt and U. Jürgens, "Acoustical correlates of affective prosody," J. Voice **21**, 531–540 (2007).

[27] J.-A. Bachorowski and M. J. Owren, "Not all laughs are alike: Voiced but not unvoiced laughter readily elicits positive affect," Soochow J. Math. **12**(3), 252–257 (2001).

[28] K. R. Scherer, "Speech and emotional states," in *Speech Evaluation in Psychiatry*, edited by J. K. Darby (Grune and Stratton, New York, 1981), pp. 189–220.

[29] T. Johnstone and K. R. Scherer, "Vocal communication of emotion," in *Handbook of Emotions*, edited by M. Lewis and J. M. Haviland-Jones (Guildford, New York, 2000), pp. 220–235.

[30] G. Fairbanks and W. Pronovost, "Vocal pitch during simulated emotion," Science **88**, 382–383 (1938).

[31] L. Anolli and R. Ciceri, "The voice of emotions: Steps to a semiosis of the vocal non-vocal communication of emotion," in *Oralité et gestualité—Interactions et comportments multimodaux dans la communication*, edited by C. Cavé, I. Guaitella, and S. Santi (L'Harmattan, Paris, 2001), pp. 175–178.

[32] R. Banse and K. R. Scherer, "Acoustic profiles in vocal emotion expression," J. Pers. Soc. Psychol. **70**, 614–636 (1996).

[33] L. Leinonen, T. Hiltunen, I. Linnankoski, and M.-L. Laakso, "Expression of emotional-motivational connotations with a one-word utterance," J. Acoust. Soc. Am. **102**, 1853–1863 (1997).

[34] G. Fairbanks and L. Hoaglin, "An experimental study of the durational characteristics of the voice during the expression of emotion," Speech Monogr. **8**, 85–90 (1941).

[35] I. Fónagy and K. Magdics, "Emotional patterns in intonation and music," Z. Phon. Sprachwissenschaft Kommunikationsforsch. **16**, 293–326 (1963).

[36] M. Bulut and S. Narayanan, "On the robustness of overall F0-only modifications to the perception of emotions in speech," J. Acoust. Soc. Am. **123**, 4547–4558 (2008).

[37] P. Boersma and D. Weenink, "Praat: Doing phonetics by computer," Ver. 4.02.04, http://www.praat.org (Last viewed July 16, 2008).

[38] D. G. Childers, *Modern Spectrum Analysis*, (IEEE, New York, 1978).

[39] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C: The Art of Scientific Computing*, (Cambridge University Press, New York, 1992).

[40] J. Hillenbrand, L. A. Getty, M. J. Clark, and K. Wheeler, “Acoustic characteristics of American English vowels,” J. Acoust. Soc. Am. **97**, 3099–3111 (1995).

[41] D. P. Szameitat, C. J. Darwin, A. J. Szameitat, D. Wildgruber, A. Sterr, S. Dietrich, and K. Alter, “Formant characteristics of human laughter,” in Interdisciplinary Workshop on the Phonetics of Laughter (Satellite Workshop of the 16th ICPHS 2007), Saarbrücken, Germany, 2007, pp. 9–13.

[42] M. Kienast and W. F. Sendlmeier, “Acoustical analysis of spectral and temporal changes in emotional speech,” in Proceedings of the ISCA Workshop on Emotion and Speech, Belfast, Northern Ireland, 2000, pp. 92–94.

AQ: #4 [43] K. R. Scherer, “Vocal affect signalling: a comparative approach,” in *Advances in the Study of Behavior*, edited by J. Rosenblatt *et al.* (Academic, New York, 1985), pp. 189–244.

[44] J. R. Davitz, “A review of research concerned with facial and vocal expression of emotion,” in *The Communication of Emotional Meaning*, edited by J. R. Davitz (McGraw-Hill, New York, 1964), pp. 13–29.

[45] J. A. Russell, “Is there a universal recognition of emotion from facial expression? A review of cross-cultural studies,” Psychol. Bull. **115**, 102–141 (1994).

[46] D. E. Mowrer, L. L. LaPointe, and J. Case, “Analysis of five acoustic correlates of laughter,” J. Nonverb. Behav. **11**, 191–199 (1987).

[47] L. L. La Pointe, D. M. Mowrer, and J. L. Case, “A comparative acoustic analysis of the laugh responses of 20- and 70-year-old males,” Int. J. Aging Hum. Dev. **31**, 1–9 (1990).

[48] R. R. Provine and Y. L. Yong, “Laughter: A stereotyped human vocalization,” Ethology **89**, 115–124 (1991).

[49] J. D. Boeke, “Mikroskopische phonogrammstudien (Microscopic phonogramm studies),” Pfluegers Arch. Gesamte Physiol. Menschen Tiere **50**, 297–318 (1891).

[50] P. A. Milford, “Perception of laughter and its acoustical properties,” Ph.D. thesis, Pennsylvania State University, Pennsylvania, 1980.

[51] C. Bickley and S. Hunnicutt, “Acoustic analysis of laughter,” in International Conference of Spoken Language Processing, 1992, pp. 927–930.

[52] M. M. Makagon, E. S. Funayama, and M. J. Owren, “An acoustic analysis of laughter produced by congenitally deaf and normally hearing college students,” J. Acoust. Soc. Am. **124**, 472–483 (2008).

[53] E. E. Nwokah, H.-C. Hsu, P. Davies, and A. Fogel, “The integration of laughter and speech in vocal communication: A dynamic systems perspective,” J. Speech Lang. Hear. Res. **42**, 880–894 (1999).

[54] J. D. Boeke, “Mikroskopische phonogrammstudien (Microscopic phonogramm studies),” Pfluegers Arch. Eur. J. Physiol. **76**, 497–516 (1899).

[55] S. Kipper and D. Todt, “Dynamic-acoustic variation causes differences in evaluations of laughter,” Percept. Mot. Skills **96**, 799–809 (2003).

[56] M. S. Edmonson, “Notes on laughter,” Anthropol. Linguist. **29**, 23–34 (1987).

[57] E. E. Nwokah, P. Davies, A. Islam, H.-C. Hsu, and A. Fogel, “Vocal affect in three-year-olds: A quantitative acoustic analysis of child laughter,” J. Acoust. Soc. Am. **94**, 3076–3090 (1993).

[58] P. Ekman, “What we have learned by measuring facial behavior,” in *What the Face Reveals*, edited by P. Ekman and E. L. Rosenberg (Oxford University Press, New York, 1997), pp. 469–485.

[59] J. A. Bea and P. C. Marijuán, “The informal patterns of laughter,” Entropy **5**, 205–213 (2003).

[60] W. Wundt, *Völkerkunde. Eine Untersuchung der Entwickungsgesetze von Sprache, Mythos und Sitte*, Die Sprache Band 1 (*Ethnology. An Investigation of the Development of Language, Myth, and Custom*, Language, Vol. **1**) (Kröner, Leipzig, 1900).

[61] F. Kainz, *Psychologie der Sprache (Psychology of Speech)* (Ferdinand Enke, Stuttgart, 1940).

[62] K. R. Scherer, “Affect bursts,” in *Emotions: Essays on Emotion Theory*, edited by S. H. M. van Goozen, N. E. van de Poll, and J. A. Sergeant (Lawrence Erlbaum, Hillsdale, NJ, 1994), pp. 161–193.

[63] C. Niemitz, “Visuelle Zeichen, Sprache und Gehirn in der Evolution des Menschen—Eine Entgegnung auf McFarland (Visual Signs, Language and the Brain in the Evolution of Humans—A Reply to McFarland),” Z. Sem. **12**, 323–336 (1990).

[64] P. Lieberman, *On the Origins of Language* (Collier Macmillan, London, 1975).

[65] U. Jürgens, “Das Trojansche Klassifikationsschema emotionaler Intonation und seine Anwendung auf Totenkopfaffenlaute (The Trojan classification scheme for emotional intonation and its application to the squirrel monkey),” Bonn. Zool. Beitr. **44**, 141–146 (1993).

[66] U. Jürgens, “Vocalization as an emotional indicator—A neuroethological study in the squirrel monkey,” Behaviour **69**, 88–117 (1979).

[67] I. Eibl-Eibesfeldt, *Ethology: The Biology of Behavior*, (Holt, Rinehart and Winston, New York, 1970).

[68] E. E. Nwokah, H.-C. Hsu, O. Dobrowolska, and A. Fogel, “The development of laughter in mother-infant communication: Timing parameters and temporal sequences,” Infant Behav. Dev. **17**, 23–35 (1994).

AUTHOR QUERIES — 025907JAS

- #1 Au: Please check definition of ″FFT″.
- #2 Au: Please supply the full journal title, the coden, and/or ISSN for the journal in refs. 5, 8, 18, 34, 35, 46, 56, and 63.
- #3 Au: PLEASE SUPPLY PAGE RANGE IN REF. 5.
- #4 Au: Please supply all editors in Ref. 43.