APRIL 14 2025

Production and perception of volitional laughter across social contexts

Virgile Daunay ; David Reby ; Gregory A. Bryant ; Katarzyna Pisanski

J. Acoust. Soc. Am. 157, 2774–2789 (2025)
https://doi.org/10.1121/10.0036388

JASA ARTICLE

Production and perception of volitional laughter across social contexts

Virgile Daunay,[1,2,a] David Reby,[1] Gregory A. Bryant,[3] and Katarzyna Pisanski[1,2]
[1]ENES Bioacoustics Research Laboratory, CRNL Center for Research in Neuroscience in Lyon, University of Saint-Étienne, 42023 Saint-Étienne, France
[2]DDL Dynamics of Language Lab, CNRS French National Centre for Scientific Research, University of Lyon 2, 69363 Lyon, France
[3]Department of Communication, Center for Behavior, Evolution, and Culture, University of California, Los Angeles, California 90095, USA

ABSTRACT:
Human nonverbal vocalizations such as laughter communicate emotion, motivation, and intent during social interactions. While differences between spontaneous and volitional laughs have been described, little is known about the communicative functions of volitional (voluntary) laughter—a complex signal used across diverse social contexts. Here, we examined whether the acoustic structure of volitional laughter encodes social contextual information recognizable by humans and computers. We asked men and women to produce volitional laughs in eight distinct social contexts ranging from positive (e.g., watching a comedy) to negative valence (e.g., embarrassment). Human listeners and machine classification algorithms accurately identified most laughter contexts above chance. However, confusion often arose within valence categories, and could be largely explained by shared acoustics. Although some acoustic features varied across social contexts, including fundamental frequency (perceived as voice pitch) and energy parameters (entropy variance, loudness, spectral centroid, and cepstral peak prominence), which also predicted listeners' recognition of laughter contexts, laughs evoked across different social contexts still often overlapped in acoustic and perceptual space. Thus, we show that volitional laughter can convey some reliable information about social context, but much of this is tied to valence, suggesting that volitional laughter is a graded rather than discrete vocal signal. https://doi.org/10.1121/10.0036388

(Received 11 September 2024; revised 19 March 2025; accepted 23 March 2025; published online 14 April 2025)

[Editor: Susanne Fuchs] Pages: 2774–2789

I. INTRODUCTION

During social interactions, people use their voices for many communicative functions (Pisanski and Bryant, 2019). We speak, laugh, cry, and scream, all in the service of social signaling (Anikin *et al.*, 2018). Yet, the ways we communicate with our voices, as in all other modalities, are profoundly shaped by social context. Even the clearest spoken utterance is often not understandable independent from the social situation in which it occurs (Sperber and Wilson, 1995). In the case of nonverbal vocalizations, such as laughter, this is also likely the case. But what information *is* present in a laugh?

Intuitively, it is easy to assume there are different kinds of laughs linked to various intentions (e.g., a melodic flirtatious laugh versus a monotonous, disinterested chuckle) and that their meanings are recognizable without context. Nevertheless, such intuitive assumptions have not been extensively tested. Here, examining volitional (voluntary) laughter in both its production and perception, we present evidence that there are some minimal distinctions in how people laugh across different social scenarios, and that these distinctions are perceptible to human listeners and machine algorithms. At the same time, we show that there is substantial ambiguity in this vocal signal, with laughter types overlapping in both acoustic and perceptual space in predictable ways that reflect their emotional valence, wherein valence can be defined as intrinsic pleasantness (positive) or unpleasantness (negative) of a situation leading to an emotion (Briefer, 2020; Goudbeek and Scherer, 2010; Kelly *et al.*, 2017). Acoustic structure is thus only one (potentially minor) piece in the puzzle of how people interpret laughter in context.

Researchers in human vocal communication have identified several categorical distinctions in laughter and other non-linguistic vocalizations. One well established distinction is between spontaneous (reflexive) versus volitional (voluntary) vocal expressions, which differ in their underlying vocal control mechanisms and in their acoustic structures (Bryant *et al.*, 2018; Bryant and Aktipis, 2014; Lavan *et al.*, 2016; Lavan *et al.*, 2017). This is largely due to the dual pathway vocal production system in humans (Ackermann *et al.*, 2014; Owren *et al.*, 2011; Scott *et al.*, 2014). An evolutionarily conserved vocal emotion system underlies the spontaneous production of our nonverbal vocal repertoire, including spontaneous laughter, crying, pain shrieks, sexual calls, and more. Yet, humans have a species-specific volitional speech system. Neural projections integrating motor cortex and language centers of the brain innervate speech articulators and laryngeal musculature,

[a]Email: virgile.daunay@cnrs.fr

https://doi.org/10.1121/10.0036388

affording volitional speech, as well as volitional emulations of our conserved non-linguistic vocal emotion repertoire, including laughter. That is, people can voluntarily laugh, cry, and scream in often convincing ways (Anikin and Lima, 2017; Bryant, 2020; Pisanski *et al.*, 2016) owing to direct connections between our brains and our vocal anatomy (Ackermann *et al.*, 2014).

We refer to these types of voluntary vocalizations as volitional because they require some degree of vocal control (Pisanski *et al.*, 2016). The remarkable control that humans have over our vocal signals is quite unique among terrestrial mammals, including primates (Ackermann *et al.*, 2014). The rarity of vocal control among primates underscores the critical need to study and understand how and why humans have evolved the capacity to produce vocalizations voluntarily, and how these volitional calls function in social life (Pisanski *et al.*, 2016). Not only did this capacity for vocal control allow our species to evolve speech, it also allows us to voluntarily modulate our non-linguistic vocalizations to communicate, exaggerate, or entirely "fake" various emotional states, such as when we embellish our pain cries (Raine *et al.*, 2018), or produce laughter entirely on demand. Few studies have examined this rare human capacity in the context of volitional laughter.

Most laughter in human social interactions is actually volitional (Mazzocconi *et al.*, 2020). We laugh during conversations for many pragmatic reasons, and we weave laughter into our conversations in fine-grained ways. Laughter is a powerful social tool for communicating subtle, often indirect meaning in the context of ostensive communication (Bryant, 2023; Sperber and Wilson, 1995). Several researchers have described the dynamics of conversational laughter, such as its timing. For example, volitional laughter tends to follow a more learned and structured rhythmic and phonemic pattern, exhibiting stable vowel sequences more frequently than spontaneous laughter, which tends to be more irregular. It typically occurs at specific points in conversation, such as between words and conversational turns, and can sometimes be incorporated into speech syllables (Provine, 1993; Vettin and Todt, 2004). More recently, Mazzocconi and colleagues (2020) described conversational laughter as having propositional content and operating in complex ways, depending on social and linguistic context, closely linked to events or states they identified as "laughables." Much earlier, Jefferson *et al.* (1987) adopted a similar approach in the tradition of conversational analysis, describing the nuanced pragmatic functions and incredible variability in how people laugh. If we are examining acoustic distinctions in laughter that map onto social communicative functions, volitional laughter is a good place to start given the highly variable social functions that it fulfills during discourse (Jefferson *et al.*, 1987; Vettin and Todt, 2004). There have been several efforts to characterize possible volitional laugh "types," but as Mazzocconi *et al.* (2020) rightly pointed out, these efforts have been hindered by confusing levels of analysis, such as creating classification schemes that mix laughter acoustics, functions, and triggers. For any proposed laugh type, all of these aspects would need to be described.

Like other non-linguistic vocalizations, laughter follows a *form* to *function* mapping (Bryant, 2020; Morton, 1977, p. 77; Owren and Rendall, 2001; Pisanski *et al.*, 2022). According to the form–function framework, the structural acoustic features of signals (forms) are often shaped by selection for their particular communicative uses (functions). Thus, we can make many straightforward predictions about which acoustic features we should expect in any given vocal signal or call type depending on what that signal is designed to accomplish. For instance, we can predict the loud and low-pitched noisy elements in an aggressive growl that evolved to enhance intimidation (Anikin *et al.*, 2024), the harsh and chaotic elements of a baby's distress cry that function to draw attention and inhibit habituation (Koutseff *et al.*, 2018; Lockhart-Bouron *et al.*, 2023), or the joyful high-pitched laughing of children playing, signaling pleasure and arousal (Nwokah *et al.*, 1993). Form–function relationships are not only relevant for distinguishing between different call types but can also apply within a single call type, such as laughter. This means that *graded* structural variations in vocalizations (i.e., "forms") can correspond to different functions. As a result, we should expect this form–-function principle to apply to any vocal behavior.

Graded vocal signaling systems are characterized by acoustic forms varying in a relatively smooth and continuous fashion, for example, along a continuum from negative to positive in valence and/or as a function of arousal and intensity, rather than forming distinct acoustic signatures or different call types (Anikin *et al.*, 2018; Briefer, 2012, 2020; Engelberg *et al.*, 2021; Lockhart-Bouron *et al.*, 2023). For instance, in human and animal vocalizations, the acoustic parameters—fundamental frequency (perceived as pitch), duration, harmonics-to-noise ratio, spectral centroid, and amplitude—have all been shown to vary gradually across emotional contexts, particularly in relation to valence and arousal (Briefer, 2012; Pinheiro *et al.*, 2021; Sauter *et al.*, 2010a; Wood *et al.*, 2017). Taken together, this gives us solid theoretical reasons to expect a fair amount of ambiguity and overlap in laughter acoustic structure that we expected to be strongly linked to valence, perhaps especially in volitional laughter, as it is under voluntary control and may maximally incorporate valence signals.

From this principle, we can expect, for example, that laughter intended to communicate amusement will be higher pitched than laughter intended to communicate malice. Importantly, we argue that form–function mappings may be predicted largely by the positive or negative intent (valence) of the laugh, which is closely linked to its communicative function, from social bonding to exclusion (Bryant, 2020; Scott *et al.*, 2014). One common strategy for categorizing laughter is to map acoustic features to emotion categories, which are also linked to form–function correspondences and social intentions (Anikin and Lima, 2017; Bryant, 2021; Cosmides, 1983; Fernald, 1992). Cowen and colleagues (2019) used machine classification methods to showcase the

https://doi.org/10.1121/10.0036388

many nuanced affective distinctions that classifiers and human judges can make in the emotional acoustic space, including laughter. Traditionally considered primarily a positive emotional vocal signal, it is now widely recognized that laughter can, in fact, accompany virtually any emotion and occur in any social context, including nervousness, anger, and sadness (Provine, 2001; Scott *et al.*, 2014).

Some researchers report acoustic distinctions across so-called laughter types. In one study, Szameitat *et al.* (2009a) had eight professional actors laugh freely in emotionally induced conditions of joy, being tickled, taunting another person, and schadenfreude (i.e., feeling pleasure from another's misfortune). The laughter recordings were then classified by trained listeners, with correctly classified laughs acoustically analyzed for frequency, intensity, and temporal properties. Discriminant analysis, using a reduced set of acoustic variables, successfully classified the four categories of laughter with 84% accuracy, with reasonable performance when using only single acoustic parameters among those selected. These data suggested that different communicative functions shape the acoustic features of laughter. For instance, joyful laughter tended to be higher pitched and faster, reflecting positive affect. Taunting laughter, conversely, was lower pitched with a higher center of gravity, associated with aggressive emotional intent (see also Szameitat *et al.*, 2022). In other work, the same researchers also found that listeners could discriminate between four distinct kinds of laughter based on emotional categories and also on the affective dimensions of arousal and valence (Szameitat *et al.*, 2009a). The chosen laughter categories in this work reflected actual social communicative intentions in real interactions, but several aspects of the study methods, in combination with one another, maximized the likelihood of finding distinctions in laughter across contexts. Namely, professional actors produced many of the laughter tokens, which underwent a selection process prior to being classified by listeners, essentially removing many ambiguous laughs in the process. Acoustic variables were then reduced to a set that were maximally distinct from one another. In the real world of actual laughter by ordinary interactants, we may not expect such clear distinctions between what are generally fairly blurred categories of social intent.

More recent work by Wood *et al.* (2017) examined whether laughter acoustics were associated with the judged social functional categories of reward, affiliation, and dominance. Volitional laughter was taken from a sound library and acoustically analyzed. The acoustic variables were then used to predict listeners' judgments of the social communicative functions of the laughs. A strength of this research is that the categories were rooted in social action and intention rather than in more abstract emotion terminology, giving the study relatively higher ecological validity. However, the laughter itself was not actually produced in the social contexts it was linked to perceptually. Other work demonstrated that laughter tokens matched in intensity and swapped out in real recorded interactions were not distinguishable by listeners, suggesting that laughter acoustics are quite underdetermined, and instead social context drives people's judgments (Curran *et al.*, 2018; Rychlowska et al., 2022).

In the current study, we aimed to address these apparent contradictions and examined (1) whether volitional laughter acoustically encodes distinct social contextual information; (2) whether human listeners and machine algorithms can accurately decode that social information, and if so; (3) which acoustic features predict laughter categorization and perception. For experimental control, we standardized our laugh contexts and eliminated social factors that could influence volitional laughter production and perception, such as relationship quality (Bryant *et al.*, 2016). Thus, we recorded people producing laughter in response to written prompts. We created eight distinct social contexts with a specific vignette for each. These contexts, partly based on previous theoretical and empirical frameworks (Devillers and Vidrascu, 2007; Lavan *et al.*, 2016; Mazzocconi *et al.*, 2020; Nikopoulos, 2017; Ruch *et al.*, 2014; Scott *et al.*, 2014; Szameitat *et al.*, 2022), are not exhaustive but include clear emotional distinctions and speaker intentions as well as interactional aspects, from the "bright" (positive) side to the "dark" (negative) side of laughter. Our eight laughter contexts included: *amused*, *colaughter, relief, acquaintance*, *mocking*, *malicious*, *sarcastic* and *nervous*. We examined production and perception differences across social contexts specifically (i.e., situational interpersonal circumstances that contextualize a communicative signal), but we conceptualize these differences largely in terms of the vocalizers' intended communicative functions. For example, in the amused social context, we are examining the interpretation of an amused laugh functioning as a signal of positive affect and affiliation. However, in the mocking social context, we are anticipating the communication of negative affect and a lack of affiliation toward an assumed listener. While we predicted at least some acoustic and perceptual distinctions in laughter across these social contexts, we also expected high overlap in laughter production and perception for laughs sharing the same (positive or negative) valence.

II. METHODS

A. Vocal recording and analysis

We audio-recorded 49 students [mean age = 22.20 years; standard deviation [SD] = 2.88; 30 females) from Jean Monnet University (Saint-Étienne, France). Vocalizers completed a brief questionnaire indicating their native language, as well as their sex, age, and height to consider anatomical traits that might influence laugh acoustics (Bachorowski *et al.*, 2001; Lavan *et al.*, 2019). All vocalizers provided informed consent, and none reported currently suffering from any conditions that might affect their voices (e.g., cold, sore throat, voice pathologies).

Laughs ($N = 382$) were recorded digitally (48 kHz/16 bit) in a sound-attenuated booth. Vocalizers stood 80 cm from a Tascam DR05 portable microphone (TEAC Corporation, Tokyo, Japan), positioned on a desk with a laptop. After recording neutral sentences and vowels,

https://doi.org/10.1121/10.0036388

participants were presented with eight specific laughter contexts in a randomized order. They were instructed to take as much time as needed to imagine themselves in each scenario, which was described in a short text vignette displayed on a monitor (see Table S1 in the supplementary material). They were then asked to voluntarily produce a laugh that they felt best reflected the given context. To preserve the ecological validity of their responses and capture a more natural volitional vocalization, participants were neither coached nor given practice trials, and only their first laughter attempt for each scenario was considered for analysis. This protocol is based on previous studies using similar methods (Kamiloğlu *et al.*, 2020; Lima *et al.*, 2013; Raine *et al.*, 2018; Raine *et al.*, 2019a, Raine *et al.*, 2019b; Sauter and Scott, 2007; Schröder, 2003).

The eight laughter contexts were chosen to encompass both positively and negatively valanced scenarios (see Table S1 in the supplementary material). Four contexts were categorized as relatively negative, including laughs emitted with sarcasm (*sarcastic*, $n=49$), during a nervous moment (*nervous*, $n=48$), to mock someone else (*mocking*, $n=48$), or the laugh of an evil character in a cartoon (*malicious*, $n=49$). The other four contexts were more positive, including polite laughter produced to signal social affiliation (*acquaintance*, $n=45$), laughter emitted while watching a comedy (*amused*, $n=49$), laughter in response to a situation of relief, such as when receiving good news (*relief*, $n=46$), and laughter shared with a friend (*colaughter*, $n=48$). After producing each laugh, speakers were asked to rate their perceived positivity of the context on a scale from extremely negative (1) to extremely positive (100). A single continuous recording of each participant was obtained and each laugh token was excised manually.

All edited laughter samples were analyzed acoustically using the *Soundgen* R package (Anikin, 2019). Acoustic measures were taken from the entire laughter bout for each laughter context, spanning from the onset of visible acoustic energy (the first burst or glottal cycle) to the offset of energy (final inspiratory element) in the final burst [Fig. 1(A)]. Pitch contours were manually verified for each laugh following an initial automatic detection using *Soundgen* (Anikin, 2019). The "analysis" and "segment" functions of this package provided measurements of numerous acoustic parameters. We focused our analyses on the 24 most relevant acoustic features (see Table S2 in the supplementary material) based on recent research on laughter and other nonverbal vocalizations that has demonstrated the communicative relevance of these acoustic

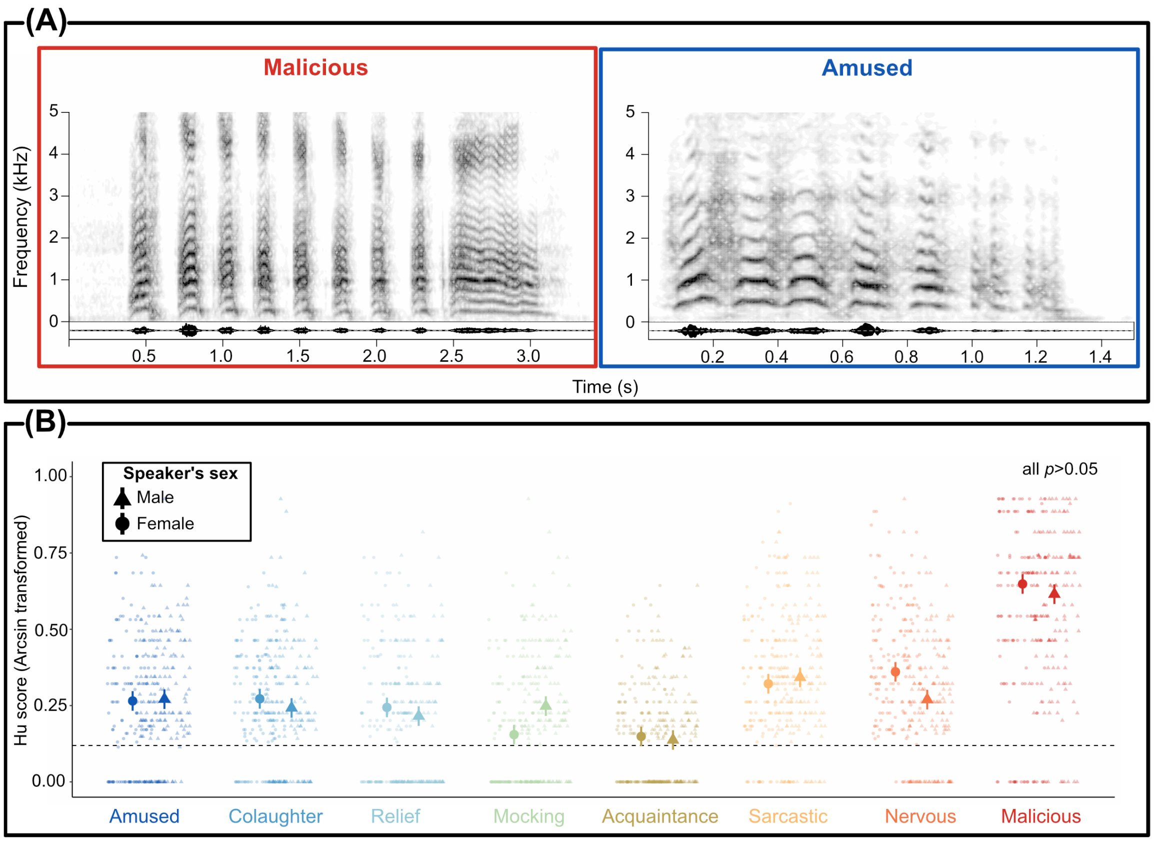

FIG. 1. Human listeners recognize laughter contexts better than chance but with low accuracy. (A) Sample waveform and narrowband spectrogram (Gaussian window, 44.1 kHz sampling rate, 0–5 kHz) of a 3 s malicious laughter bout with nine bursts and a 1.5 s amused laughter bout with seven bursts produced by the same woman. In these examples, malicious laughter is lower pitched and more regular than is amused laughter. (B) Hu scores (arcsine-transformed unbiased hit rates) indicating listener's ability to discriminate laughter contexts above the arcsine-transformed chance level (0.12, dashed line), controlling for individual perceptual biases. Each context was significantly recognized above this corrected chance level [$t_{(1780)} > 1.9$; $p < 0.05$]; only the acquaintance context showed a negligible effect size (Cohen's $d = 0.09$). However, note that except for malicious laughter, Hu scores were not higher than 30, showing only a moderate capacity in human listeners to recognize most laughter contexts.

features (e.g., Bachorowski and Owren, 2001; Briefer, 2012; Bryant and Aktipis, 2014; Pinheiro *et al.*, 2021; Sauter *et al.*, 2010a).

The duration of each laugh was measured based on the length of each bout. For all other acoustic parameters, we measured the mean, including fundamental frequency (f_o), amplitude (amp), and the frequency of maximum amplitude (peakFreq). Parameters describing voice quality were also measured: how loud the laugh sounds (an estimate of subjective loudness in sones), how rough it sounds (roughness based on the proportion of energy within the range of perceptible temporal modulation frequencies), and how noisy it is based on the harmonics-to-noise ratio (HNR), Wiener entropy, and cepstral peak prominence (CPP), which also measure vocal instability. The regularity of laughter was measured by the mean frequency of frequency modulation (vibrato, jitter), whereas the center of gravity of each spectrum (SpecCentroid, the center of gravity of spectrum) indicated the relative influence of lower versus higher frequencies indicating how "bright" a laugh sounds. Temporal parameters included the duration between laugh bursts [mean intervoicing interval (IVI)], the duration of these interburst intervals, the number of bursts (i.e., continuous segment different from noise) per second (nburst.s-1), and the rate of the intervoicing interval as defined by Bryant and Aktipis (2014) to measure the average rate of unvoiced segments per call in each laugh. Finally, we measured the proportion of voiced parts across all laughter bouts (voicing) and within calls (voiced_noSilence). For all these variables, we calculated the coefficient of variation (CV), which expresses relative variability by calculating the ratio between the standard deviation and the mean of the given acoustic parameter. This approach allows for meaningful comparisons across individuals with different mean values, minimizing the imbalance due to these absolute individual differences.

B. Perception experiment

We created a forced-choice experiment using the Labvanced platform (Scicovery GmbH, Paderborn, Germany) (Goeke *et al.*, 2017). Listeners were recruited online through Prolific (Prolefic Academic Ltd., London, UK) and compensated at the recommended rate of £8/h (∼$10.30/h) for their time. At the beginning of the experiment, we collected demographic information from listeners, including age and gender (see S1 in the supplementary material). We also collected information about musical experience (e.g., whether participants play a musical instrument) as well as vocal performance experience, such as acting or singing. Participants reported whether they have such experience and its duration in years (*duration of musical experience* and *duration of vocal performance experience*). We instructed listeners to complete the experiment in a quiet environment.

The sample size of listeners was determined through a power analysis using the (Champely, 2020) R package for a z-proportion test given $d = 0.2$, $power = 0.80$, and $a = 0.05$. To detect a small effect size on this task, 155 listeners were required. To ensure listeners were wearing headphones, we incorporated a verification task (Woods *et al.*, 2017) and automatically excluded listeners who failed. Additionally, two random attention checks were implemented in which listeners were asked to indicate when an artificial tone was played instead of a laughter stimulus. Before the actual experiment, listeners were presented with a sound file containing two volitional laughs (one loud and one quiet) not included in the experiment. They were instructed to raise their computer volume until they could clearly hear the quiet stimulus without feeling discomfort from the louder one. Subsequently, listeners were asked to not adjust their volume during the whole experiment, and this was reconfirmed at the end of the experiment. Participants who reported modifying the volume or not wearing headphones were removed from the analysis. Thus, from 168 English-speaking listeners, 12 were excluded for a final sample size of 156 listeners (mean age = 29.8 years, range: 19–69 years; 76 females).

The perception experiment began by presenting each laugh context and the associated vignette used during the production task (see Table S1 in the supplementary material). Listeners were first required to rate the perceived emotional valence of each context on a 100-point scale, ranging from extremely negative to extremely positive, following the same protocol as for the speakers. Thus, in this study, we collected valence ratings of laughter contexts from both the speakers and the listeners.

In a second step, 80 volitional laughs (ten for each context, with an even distribution of male and female laughs) were randomly selected from the 382 recordings and presented to listeners in random order. Their task was to select which context the laugh was produced in, from among the eight different contexts presented in a random order. The same vignette used for the vocalizer task, initially presented in full at the start of the experiment, was provided in a shortened version during each trial as a reminder. Listeners could replay laughs an unlimited number of times before selecting the context.

C. Statistical analysis

All statistical analyses were conducted using *R* software (R Core Team, 2021). To first test whether ratings of emotional valence varied by laughter context as predicted, we ran a linear mixed model on vocalizer valence ratings for each context, with vocalizer ID as a random effect. We carried out a second similar analysis on the listeners' valence ratings of the laughter contexts and associated vignettes. We ran a third linear mixed model to investigate any differences between vocalizer and listener valence ratings.

We then used machine classification algorithms in conjunction with human classification to test how laughter varies by context based on a set of acoustic parameters (focusing on the production aspect) and to assess listeners' ability to identify these contexts (focusing on the perception aspect). If the accuracy of listeners' classifications is higher than the accuracy of the classification algorithms, it likely suggests that the selected acoustic features do not capture the full range of acoustic variation among contexts. Conversely, low accuracy in listeners' classifications could indicate that listeners are unable to detect all the acoustic variation present. Last, if the results from both methods (machine and human) are similar, this would suggest that the chosen acoustic features in our analysis are similar to those features tracked by listeners.

1. Classification algorithms

Choosing a classification algorithm often involves a trade-off between several important features including the structure of the data, the computation time, interpretability, and robustness. Different classifiers will also exploit the data in different ways (Arnaud *et al.*, 2023). For this study, we analyzed two supervised classifiers: discriminant function analysis (DFA) and random forest (RF). By employing two distinct classifiers, we aimed to balance robustness and interpretability while leveraging each method's specific strengths to explore the relative roles of acoustic variables in classifying laughter contexts (additional details regarding the models can be found in S2 in the supplementary material).

These models were evaluated across 1000 iterations, wherein each iteration involved generating a new pair of training (80% of the data) and test (20% of the data) sets, with stratification based on laughter context to ensure class balance. The acoustic variables were used directly without dimensionality reduction to maximize interpretability and assess the relative contributions of each feature (Anikin and Persson, 2017; Wadewitz *et al.*, 2015). To address multicollinearity, variables exceeding a variance inflation factor (VIF) threshold of five were excluded (James *et al.*, 2013). Specifically, the following variables were removed: coefficient of variation of amplitude (VIF = 10.7), f_o bandwidth (VIF = 9.2), spectral centroid mean (VIF = 9.08), and voiced proportion (VIF = 5.1). Acoustic variables were normalized (mean = 0; SD = 1), and frequency-related measures were log-transformed before normalization. One laugh, produced in a mocking context by a female, was completely unvoiced and thus excluded from the classification analysis.

The DFA model was trained using tenfold cross-validation repeated ten times. Model performance was evaluated using accuracy, precision, recall, and $F1$-score, calculated based on predictions made for the test samples. Confusion matrices were also generated, and variable contributions were assessed using loadings. Performance metrics were averaged over 1000 iterations, with confidence intervals calculated.

Prior to model evaluation, hyperparameters of the RF were optimized using a ten-time repeated tenfold cross-validation. Initial tuning was performed via a random grid search, followed by refinement with a regular grid of five levels. The selected hyperparameters included *mtry* (number of predictors per split) = 10, *min_n* (minimum observations per node) = 20, and *n_trees* (number of trees) = 1000. The model was trained and tested using a DFA "data splitting" approach, with performance metrics calculated for each test sample. Variable importance was assessed using permutation-based importance with the *Variable Importance Plots* (*vip*) R package (Greenwell and Boehmke, 2020), which evaluates the impact of shuffling each variable on model performance. As with the DFA, metrics were averaged over 1000 iterations and confidence intervals were calculated.

2. Models examining listeners' laughter judgments

We constructed a generalized linear mixed model (GLMM) based on the binomial family (response correct or incorrect) to analyze whether listeners could reliably recognize laughter contexts based only on the acoustic signal. Generalized linear mixed models (GLMMs) allowed us to consider the random effects of the variability of responses among different listeners and the variability in laughter acoustics across speakers. The chance level was set to 0.125 (eight contexts). We defined a logit for this reference probability, which was entered into the model as an offset term. We first built an omnibus model including both vocalizers' and listeners' individual difference measures (e.g., sex, age, and musical experience) that have been shown to potentially affect vocal perception (Amorim *et al.*, 2021; Cartei *et al.*, 2020). The total time taken by the listener to complete the experiment was also considered.

The importance of these fixed effects was then estimated by the different combinations of fixed effects and their interactions using *AICc* comparisons. The best model to explain context recognition accuracy included the fixed effect of laughter context, vocalizer sex, and listener age (see Table S3 in the supplementary material). Interactions between age and context, and sex of the speaker and context, were also included. *Post hoc* analyses with z-tests (Holm correction for multiple testing) were then performed to assess the effects of each fixed factor on accuracy. For each listener, we computed the unbiased hit rate (Hu scores) (Wagner, 1993) to control for individual biases in choosing some contexts more than others. These values were arcsine-transformed and included in a second GLMM model with the same fixed and random effects. Finally, we ran *post hoc* tests with holm correction for multiple comparisons and compared arcsine-transformed Hu scores to arcsine-transformed corrected chance levels for each laughter context (Wagner, 1993).

3. Models examining laughter acoustics

We performed independent linear mixed models on each acoustic feature selected as the response variable to test how laughs varied acoustically across contexts. Models included context and vocalizer sex as fixed effects, and vocalizer identity as a random effect. To test the influence of context on each parameter, we compared each model to the null model with a likelihood ratio test. The importance of fixed effects was also considered by ranking the model using the corrected Akaike information criteria (*AICc*). We calculated marginal R^2 values to compute the proportion of variance explained by each acoustic parameter. Finally, we calculated estimated marginal means based on the fitted results of each model. On these, pairwise comparisons were made with *post hoc* tests (Holm correction for multiple testing).

We then tested the predictions of each acoustic feature on listeners' recognition across contexts through multiple generalized mixed models based on a binomial distribution (correct vs incorrect). As using all 24 acoustic parameters in one model would certainly overfit it, we built multiple generalized mixed models with one acoustic parameter and its interaction with laugh context. We included speaker and listener IDs as random effects. We examined both the *AICc*, which balances model fit and complexity and the marginal R^2, which quantifies the variance explained by fixed effects. Together, these metrics provide a comprehensive assessment of each acoustic parameter's contribution to laughter perception. Tukey's *post hoc* analyses with Holm correction were then performed to test the effects of each acoustic parameter on listeners' recognition of laughter context.

III. RESULTS

A. Context-specific laughs are recognized by both humans and machines, but not with precision

Overall, our results show that laughter acoustics can encode contextual information that is detectable by both machine algorithms and human listeners, but not with high precision, across discrete laughter contexts (Table S4). Indeed, both human listeners and classification algorithms could significantly identify the context of laughter above chance level (12.5%), but accuracy varied considerably depending on the laughter context (Table I).

For both computer algorithms and human listeners, *nervous*, *colaughter*, and *amused* contexts were consistently recognized at levels exceeding chance (from 26% for colaughter to 43% for amused laughter, RF) (see Table I). Similarly, the *malicious* context showed the highest recognition accuracy in all classification tests, exceeding 50% accuracy. In contrast, *relief*, *acquaintance*, and *mocking* laughter contexts were consistently recognized at rates not exceeding chance (<12.5%) or at low accuracy for the DFA (21% for *mocking*) (Table I).

Our models comparing human and machine performance showed that machines and humans performed highly similarly. Only minor differences were observed, notably the *malicious* context was less effectively recognized by humans, whereas *sarcasm* was more effectively recognized by humans, compared to machines (Table S5 in the supplementary material).

To control for potential biases in judging laughter contexts, we examined unbiased hit rates (Hu scores), which revealed that all eight contexts were significantly recognized higher than corrected chance levels by human listeners [Fig. 1(B)]. Vocalizer sex also predicted variance in listeners' unbiased hit rates [Fig. 1(B)], wherein *mocking* laughter [$t_{(2362)} = -4.161$; $p < 0.001$; $d = -0.47$] was significantly better recognized when the vocalizer was male. Conversely, *nervous* laughter was significantly better recognized when the vocalizer was female [$t_{(2362)} = 4.05$; $p < 0.001$; $d = 0.46$]. As with raw hit rates, we observed a slight decrease in accuracy with listener age with small effects for *malicious* [$t_{(1832)} = -5.98$; $p < 0.001$; $d = -0.24$], and negligible effects for *mocking* [$t_{(1832)} = -2.64$; $p = 0.008$; $d = -0.11$], *sarcastic* [$t_{(1832)} = -2.11$; $p = 0.035$; $d = -0.08$], and *colaughter* [$t_{(1832)} = -2.00$; $p = 0.046$; $d = -0.08$] contexts. Similar effects of vocalizer sex (Table S6 in the supplementary material) and listener age (Table S7 in the supplementary material) were found in models based on raw ratings.

In our sample of listeners, 41 self-reported as musicians and 32 reported having vocal performance experience. However, models including musical ($AICc = 13\,956$; $\chi^2_8 = 6.24$; $p = 0.62$) and vocal experience of listeners ($AICc = 13\,956$; $\chi^2_8 = 8.18$; $p = 0.41$) were no different from the null model ($AICc = 13\,947$). Similarly, the model including the total time taken to complete the experiment was not significantly different from the null model ($AICc = 13\,949$; $\chi^2_8 = 13.33$; $p = 0.10$). Thus, these variables were shown to have no significant effects on context discrimination accuracy.

B. Laughter contexts are systematically confused with those that share the same valence

Having established that social context is encoded in laughter sufficiently for both machine algorithms and human listeners to recognize contexts better than chance

TABLE I. Laughter context recognition accuracy by human listeners and machine classification algorithms based on 24 acoustic parameters. Recognition rates in bold are significantly higher than chance (12.5%).

	Human listeners		DFA		Random forest	
Laughter context	Mean	95 % CI	Mean	95 % CI	Mean	95 % CI
Malicious	**0.50**[a]	**[0.46–0.53]**	**0.63**[a]	**[0.62–0.64]**	**0.70**[a]	**[0.69–0.70]**
Nervous	**0.36**[a]	**[0.34–0.39]**	**0.27**[a]	**[0.27–0.29]**	**0.37**[a]	**[0.36–0.38]**
Sarcastic	**0.35**[a]	**[0.33–0.38]**	**0.17**[a]	**[0.16–0.18]**	**0.12**[a]	**[0.11–0.12]**
Acquaintance	0.15	[0.13–0.17]	0.14	[0.1 3–0.14]	0.07	[0.06–0.07]
Mocking	0.18	[0.16–0.2]	**0.21**[a]	**[0.20–0.22]**	0.10	[0.09–0.10]
Relief	0.17	[0.15–0.19]	0.10	[0.09–0.11]	0.01	[0.01–0.01]
Colaughter	**0.28**[a]	**[0.26–0.31]**	**0.27**[a]	**[0.26–0.28]**	**0.26**[a]	**[0.25–0.26]**
Amused	**0.28**[a]	**[0.25–0.3]**	**0.30**[a]	**[0.29–0.31]**	**0.43**[a]	**[0.42–0.44]**
Overall	**0.28**[a]	**[0.27–0.29]**	**0.28**[a]	**[0.28–0.29]**	**0.29**[a]	**[0.29–0.29]**

[a] p value < 0.001.

JASA
https://doi.org/10.1121/10.0036388

(though with low discrete classification), we tested specific confusion patterns. When listeners and machines are wrong, which laughter contexts do they most often confuse?

The confusion matrix of the two machine classification algorithms (Fig. 2) revealed that *amused* and *colaughter* were highly confused, usually more than 20% of the time. This is further illustrated by the tree plots in Fig. 2, showing a very close Euclidean distance between these two types of positive laughter. *Mocking* and *relief* laughs were also often confused with *amused* and *colaughter* laughs (>17%). These contexts formed one cluster according to our hierarchical cluster analysis. Another cluster was composed of *nervous*, *sarcastic*, and *acquaintance* laughs. These contexts were also often confused with each other (>10%); for instance, *sarcastic* laughter was misclassified in more than 23% of cases as *nervous*, whereas *nervous* laughter was misclassified in more than 16% of cases as *sarcastic*.

We found similar clusters based on the classification matrix and Euclidean distances for human listeners (Fig. 2): *Amused* and *colaughter* laughs were again often confused with one another (>25%). *Mocking* laughs were also regularly confused with *amused* and *colaughter* contexts by humans (>19%). *Nervous* laughter was often confused with *acquaintance* laughter (19%), *sarcastic* laughter (13%), or *relief* laughter (10%). Finally, *malicious* laughter was often considered *sarcastic* (20%) or as *mocking* (15%), explaining its lower recognition accuracy for humans compared to computer algorithms (Table S5 in the supplementary material).

We found that laughter contexts could be clustered based on confusions made by humans and machines, raising the possibility that confusion clusters could be explained by valence. As illustrated in Fig. 3(A), vocalizers perceived the emotional valence of their laughter contexts in a graded manner [Fig. 3(A)] [linear mixed model (LMM):

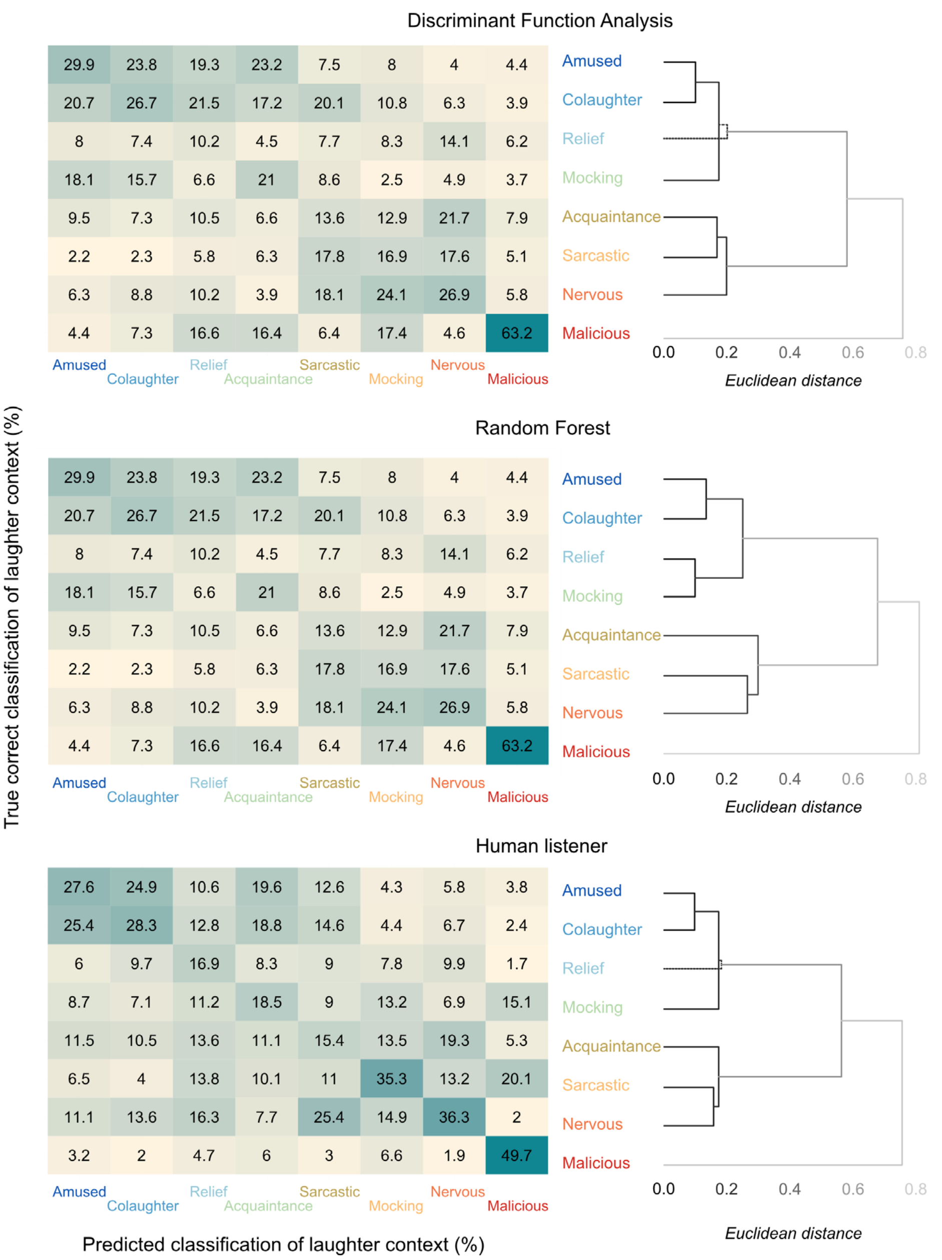

FIG. 2. Laughter contexts are systematically confused with those that share the same positive or negative valence. Heatmaps of confusion matrices (%) for context classification accuracy by the DFA, RF model, and human listeners. The *x* axes represent the true classification and the *y* axes represent the predicted classification. The color of each cell represents the proportion of contexts accurately predicted, with lighter shades showing lower classification accuracy and darker shades showing higher classification accuracy. The dendrograms on the right were obtained through hierarchical clustering using Ward D2 methods and build on Euclidean distances calculated from these confusion matrices. The gray gradients correspond to the Euclidean distances between each context. For the DFA and RF classification models, data were obtained from 1000 iterations with different train and test sets for each. For human listeners, data were derived from the perception experiment on 156 listeners. These tree plots show that some contexts are often confused with others, such as amused and colaughter.

$AICc = 3404.7$; $\chi_7 = 351.9$; $p < 0.001$]. *Post hoc* comparisons revealed that the contexts of *colaughter* [$M = 87.3$; standard error $(SE) = 3.01$], *amusement* [*marginal mean* $(M) = 85.8$; $SE = 2.98$], and *relief* ($M = 81.9$; $SE = 3.07$), and were rated similarly positively by vocalizers ($p > 0.05$) and were significantly more positive in valence compared to all other contexts ($p < 0.001$). The contexts of *mocking* ($M = 53.0$; $SE = 3.01$), *malicious* ($M = 45.2$; $SE = 3.028.4$), and *acquaintance* ($M = 42.9$; $SE = 3.1$) were rated as intermediate in valence and did not significantly differ from one another. These three contexts were rated significantly more positively ($p < 0.001$) than the *sarcastic* ($M = 29.1$; $SE = 3.0$) and *nervous* ($M = 20.1$; $SE = 3.01$) contexts, which received the most negative valence ratings by vocalizers.

Similar results were found when listeners rated the valence of laughter contexts before the listening experiment [Fig. 3(B)] (LMM: $AICc = 10\,867$; $\chi_7 = 1587$; $p < 0.001$). *Amusement* ($M = 92.5$; $SE = 1.52$), *relief* ($M = 87.8$; $SE = 1.52$), and *colaughter* ($M = 87.2$; SE $= 1.52$) were rated similarly positively by listeners ($p > 0.05$). *Acquaintance* laughter ($M = 50.9$; $SE = 1.52$) was also rated as intermediate in valence and significantly more positive ($p < 0.001$) than the *sarcastic* ($M = 29.9$; $SE = 1.52$), *nervous* laughter ($M = 29.0$l; $SE = 1.52$) and, contrary to the vocalizer's rating, the *malicious* context ($M = 25.6$; $SE = 1.52$). These laughter judgments were significantly different from one another. Finally, *mocking* laughter ($M = 15.3$; $SE = 1.52$) received the most negative valence rating from listeners ($p < 0.001$). Indeed, concerning the third linear mixed model ($AICc = 14\,292$; $\chi_7 = 1942.9$; $p < 0.001$), *malicious* and *mocking* laughter were rated significantly more negatively by listeners than vocalizers ($p < 0.001$). However, there were no significant differences

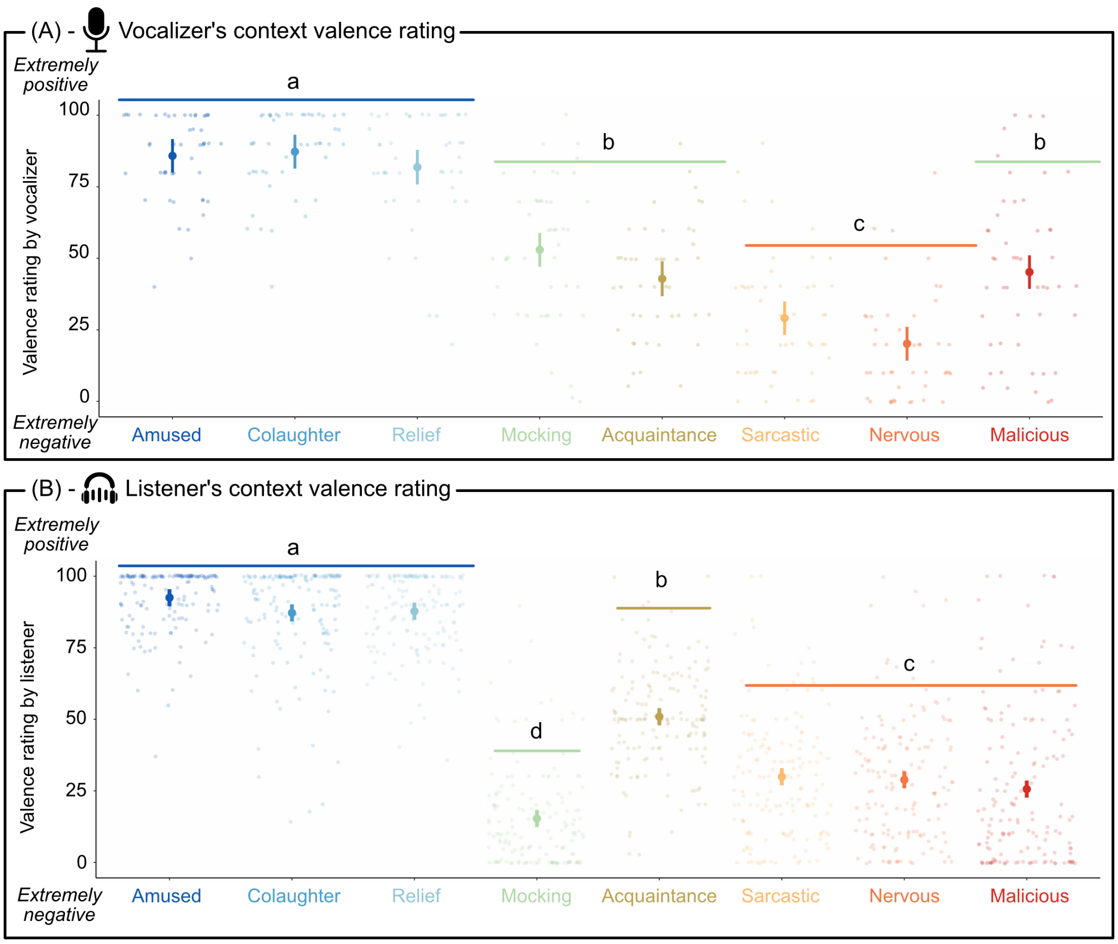

FIG. 3. Vocalizers and listeners rate the emotional valence of laughter contexts in a similar way. (A) Vocalizers' valence ratings of each laughter context after acting out the laugh. Pairwise comparisons made on estimated marginal means obtained from a linear mixed model, with individuals as random effects, significantly discriminated three levels of valence for the laughter contexts (Tukey's post hoc tests $P < 0.001$), with group "a" representing the most positively rated contexts (amused, colaughter, relief), group "b" representing significantly less positive contexts (mocking, malicious, and acquaintance), and group "c" representing the most significantly negatively rated contexts (nervous and sarcastic). (B) Listener's valence ratings of laughter contexts before the listening experiment. Pairwise comparisons made on estimated marginal means obtained from a linear mixed model with individuals as random effect significantly discriminated four levels of valence for the laughter contexts (Tukey's post hoc tests $P < 0.001$), with group 'a' representing the most positively rated contexts (amused, colaughter, relief), groups "b" and "c" representing significantly less positive contexts (acquaintance followed by malicious, nervous, and sarcastic), and finally mocking laughter as the most negatively rated context (group "d"). Overall, our results show that the perceived valence of laughter contexts was similar for vocalizers and listeners.

in valence between the other laughter ($p > 0.05$), indicating that vocalizers and listeners showed high agreement about which laughs were positive and which were negative.

In sum, as illustrated in Figs. 2 and 3, when listeners incorrectly judged the context of a laugh, the confusion was generally with laughter contexts of a similar valence (positive or negative), especially when we consider the valence ratings of vocalizers.

C. Fundamental frequency, acoustic energy, and temporal regularities differentiate context-specific laughs

We have shown that social contextual information in volitional laughter is detectable by human listeners and machine classification, and that confusions among contexts are predictable based on vocalizer's valence ratings. Next, we examined which acoustic parameters varied as a function of context (production), and how those features predicted listeners' contextual judgments (perception).

Independent linear mixed models (LMMs), with vocalizer sex and laughter context as fixed factors, showed significant variation in acoustic parameters across laughter contexts. Figure 4 illustrates forest plots of the ten acoustic features (out of all 24) that explained the most variance in laughter acoustics according to independent LMMs and marginal R^2 by social context of laughter [Fig. 4(A)]. According to the marginal R^2, the following acoustic variables explained a substantial portion of the variance in

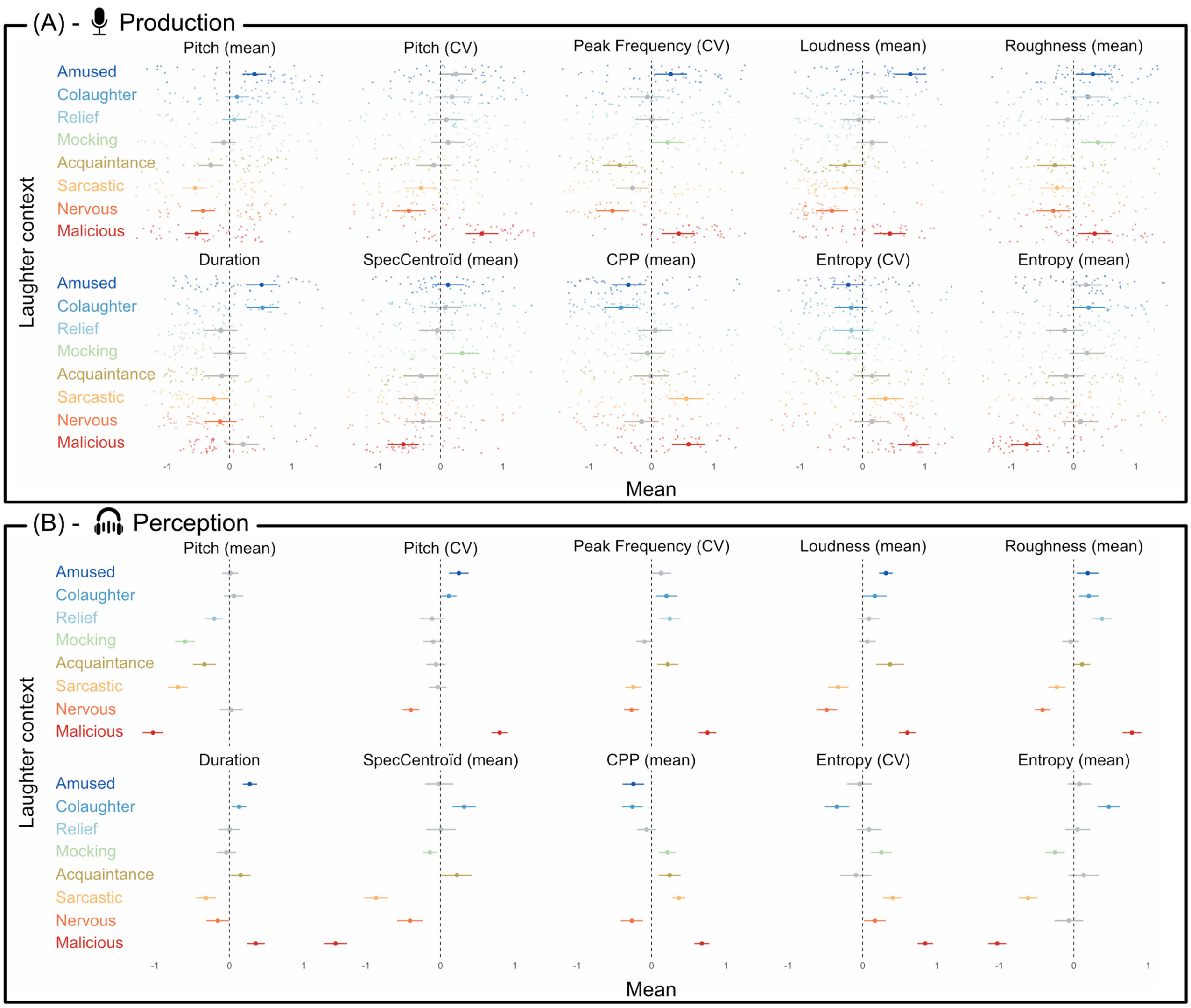

FIG. 4. Laughter acoustic structure varies across contexts, and is used by listeners to assess laughter context. (A) Forest plots of the ten acoustic features that explained the most variance in laughter acoustics according to independent LMMs and marginal R^2 by social context of laughter. Solid central markers represent scaled estimated marginal means from LMMs with 95% confidence intervals, with translucent markers showing the full distribution of the data. Colored central markers represent significant differences in marginal means compared to the average mean in Tukey's post hoc tests with Holm correction for multiple comparisons, whereas greyed-out points are not significantly different from the average mean of all contexts combined (dashed vertical line = 0). (B) The ten acoustic features that best explained variance in laughter acoustics, based on independent LMMs and marginal R^2 values, predicting listeners' recognition of each laughter context: standardized parameters and 95% confidence intervals (error bars) obtained from a GLMM based on a binomial distribution. Colored markers represent significant parameters based on a Wald approximation, and greyed-out markers show non-significant effects relative to the null effect shown by the dashed line (0). CPP, cepstral peak prominence; CV, coefficient of variation; HNR, harmonic-to-noise ratio; PeakFreq; frequency of maximum amplitude; SpecCentroid: spectral centroid.

https://doi.org/10.1121/10.0036388

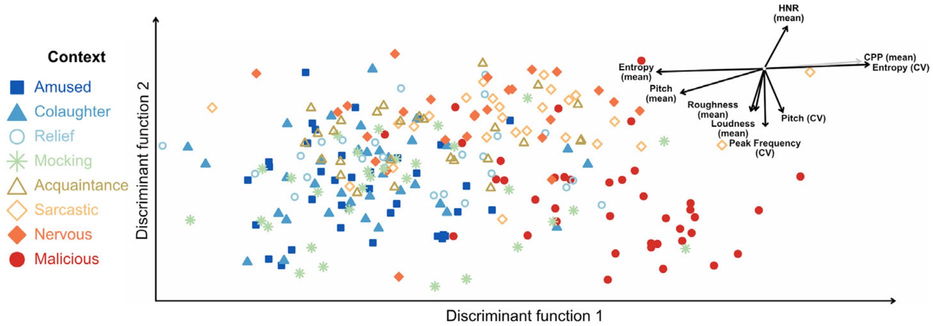

FIG. 5. Laughter acoustics vary across contexts in a graded manner that is related to emotional valence. Discriminant function analyses illustrating acoustic separation of laughs by context and valence for all vocalizers. Fundamental frequency parameters (mean and coefficient of variation of the pitch) and energy parameters (mean and coefficient of variation of entropy, mean peak frequency, subjective loudness, roughness, harmonics-to-noise ratio, and cepstral peak prominence) were the main acoustic parameters that explained variance in laughter across contexts. Each data point represents the individual context-specific laugh stimulus as a function of the first two discriminant variables that maximize individual separation, whereas colors represent the relative valence of the laughter. Note that positively valanced laughs (in blue) such as amused and colaughter cluster more to the left, whereas negatively valanced laughs (in red), such as malicious and nervous, tend to cluster to the right. The radar plot represents the loadings of the acoustic variables onto the first two discriminant functions.

laughter acoustics across fixed effects: mean f_o (56%), subjective loudness (23%), duration (20%), peak frequency (20%), and variation of entropy (19%). Similar results were found with LMMs for each sex with only laughter context as a fixed factor (Table S2 in the supplementary material). Notably, mean f_o explains similar variance in both male (18%) and female (21%) vocalizers' productions of laughter across different contexts (see Table S2 in the supplementary material).

In *post hoc* tests comparing marginal means of parameters to the average across all contexts (Fig. 4), we found that positive laughs produced in *amused*, *colaughter*, and *relief* contexts were significantly higher pitched with less entropy variation than were other laughs. *Amused* and *colaughter* laughs were also shorter and more irregular and noisier according to CPP measures. *Amused* laughs were particularly loud, with peak amplitude at higher frequencies. *Mocking* laughter showed amplitude peaks at higher frequencies and low entropy variation. *Sarcastic*, *acquaintance*, and *nervous* laughs were relatively quiet, less rough, and had peak amplitudes at lower frequencies. *Sarcastic* and *nervous* laughs were also lower in f_o, longer, and their amplitude peaks were lower. *Malicious* laughs had a lower and more variable f_o, with amplitude peaks at higher frequencies, more variability in entropy, and were louder than most other laughs (see Table S8 in the supplementary material).

Taken together, much of this acoustic variance follows predictable form–function mappings. *Malicious* laughter elicited the most extreme acoustic features, corroborating its strong perceptual salience as revealed by the perception experiment. Generally, f_o (perceived as pitch) and variation of entropy (affecting the consistency of how noisy the whole laugh sounds) were highly comparable for laughs emitted in contexts with a similar valence [see Fig. 4(A)]. Positive laughs were, as predicted by our form–function framework, relatively higher pitched than were negative laughs.

Next, we examined which acoustic features were most important for the classification algorithms when discriminating laughter contexts. While the predictive capacity of the classification algorithms is limited and this may not fully represent the absolute acoustic differences across contexts, our results showed that the same acoustic parameters that varied by laughter context based on our acoustic analyses above also predicted the performance of our classifiers. In the DFA (Fig. 5), mean entropy ($r = -0.65$) and mean pitch f_o ($r = -0.51$) both loaded negatively onto the first dimension, whereas entropy variation ($r = 0.64$) and mean CPP ($r = 0.58$) loaded positively. Mean peak frequency ($r = -0.71$), pitch variance ($r = -0.54$), loudness ($r = -0.53$), and roughness ($r = -0.52$) all negatively loaded onto the second dimension, whereas HNR loaded positively ($r = 0.52$). In the RF models, peak frequency and pitch mean were the most important acoustic parameters for discriminating laughter contexts (Figure S1).

Together, these results show that from our selection of acoustics parameters, those that most strongly differentiated laughter contexts were f_o (pitch), the distribution of energy in the laugh, as well as the regularity of the laughter bouts. Additionally, our results show that laughter contexts that are most often confused with each other have similar acoustics profiles.

D. Acoustic parameters predict how effectively listeners can judge the meaning of laughter

Using generalized LMMs, we investigated whether the acoustic parameters used by listeners to accurately assess laughter contexts were the same as those that differentiated laughs by context [Fig. 4(B)]. We found that all models were better fitted than the null model (Table S9 in the supplementary material), [$\chi^2(8) > 50$; $p < 0.001$]. However, each acoustic parameter and its interaction with context showed only weak explanatory power (*marginal* $R^2 < 0.1$). The mean spectral centroid (*marginal* $R^2 = 0.070$; $AICc = 11\,651$), mean pitch (*marginal* $R^2 = 0.060$; $AICc = 14\,331$), entropy mean (*marginal* $R^2 = 0.058$; $AICc = 11\,696$), entropy variation (*marginal* $R^2 = 0.050$; $AICc = 11\,854$), and variation of HNR (*marginal*

$R^2 = 0.041$; $AICc = 14\,469$) were the five best-fitted models with the most power to explain accuracy in listeners' judgments of laughter context.

Post hoc analyses on the effects of the acoustic parameters [Fig. 4(B)] within each context revealed that with a non-negligible effect, *malicious* contexts were better recognized when laughs were less bright (lower spectral centroid, $z = -17.73$; $p < 0.001$; $d = -0.35$), less dysphonic (higher CPP: $z = 13.21$; $p < 0.001$; $d = 0.2$), less noisy (lower mean entropy, $z = -16.85$; $p < 0.001$; $d = -0.33$) but also louder (subjective loudness, $z = 10.42$; $p < 0.001$; $d = -0.21$) with higher perceived amplitude variation (more rough: $z = 121.00$; $p < 0.001$; $d = 0.24$), higher entropy variation ($z = 15.89$; $p < 0.001$; $d = 0.31$), and with amplitude peaks at higher frequencies ($z = 12.31$; $p < 0.001$; $d = 0.24$). The pitch was also lower ($z = -14.33$; $p < 0.001$; $d = -0.28$) and more variable ($z = 14.29$; $p < 0.0011$; $d = 0.28$) in laughs that were correctly judged by listeners as *malicious*. Similarly, *sarcastic* laughs were more often correctly recognized as such when they were relatively lower in pitch ($z = -10.35$; $p < 0.001$; $d = -0.20$) and less bright (lower spectral centroid $z = -10.47$; $p < 0.001$; $d = -0.21$). Other acoustics parameters showed significant effects on context recognition, but the effects were negligible (Table S10 in the supplementary material, $d < 0.20$).

Together, these final results show that voice pitch dimensions (i.e., f_o and its variation) and acoustic measures of energy (e.g., entropy, roughness, and CPP) encode social contextual information in laughter that also predicts listeners' accuracy in decoding the meaning of laughter.

IV. DISCUSSION

Laughter is ubiquitous in social interaction and fulfills complex pragmatic functions that are not yet well understood. Some research suggests that people produce distinct laugh types associated with their communicative intentions, but other work demonstrates clear ambiguities in the acoustic structure of laughter that make laughs difficult to understand out of context. Overall, we confirmed that volitional laughter can in fact encode some social contextual information, but it is limited. Voice pitch and vocal energy parameters differentiated distinct volitional laughter contexts to some extent. Classification algorithms and human listeners were also able to correctly identify laughter contexts, though not with high precision, and relied on the same acoustic parameters. These findings support earlier work showing that laughter in specific contexts can have distinguishable features. However, we also found that classification algorithms, as well as human listeners, made predictable errors based on the emotional valence of laughter contexts. Positive contexts were often confused with other positive contexts, whereas negative contexts were often confused with other negative contexts. For instance, amused and colaughter laughs were often confused with one another, but rarely misidentified as sarcastic or nervous. Laughs of a similar valence also shared similar acoustic profiles, and this further explained listeners' (mis)assessments of laughter context. While the listeners in our study were generally able to judge the meaning of laughs, their performance was not consistently accurate, with half of the social contexts being recognized with only moderate accuracy or even below chance level. The most recognizable context, malicious laughter, was correctly identified around 50% of the time or more, substantially better than the chance level of 12.5%. We did not explicitly instruct vocalizers or listeners to adopt either the perspective of speakers or listeners when rating contexts, and this could also be a source of variation in judgments between them.

Our results are consistent with most prior work, but some differences emerged. For example, we found overall lower automated classification accuracy than reported in other research (e.g., Szameitat *et al.*, 2009b) albeit with similar judgment accuracy by human listeners (Szameitat and Szameitat, 2024), which is fairly low overall. If the notion of acoustically distinct laugh types was correct, we should not find this pattern across multiple studies. Wood *et al.* (2017) used social functional categories that map loosely onto some of the categories used in the current study. Like Wood *et al.*, we found that *mocking* and *amused* laughs, similar to their categories of dominance and reward, manifested in quite different ways, especially in vocal pitch. Wood and colleagues found that *dominance* laughs (like our *mocking* laughter) tended to be lower in f_o (pitch) and higher in their center of gravity, acoustic features that are often associated with high arousal and a larger perceived body size (Briefer, 2020; Charlton and Reby, 2016; Pisanski and Reby, 2021). Conversely, *reward* laughs (like our *amused* laughter) tended to be higher in f_o and noisier (high roughness and lower harmonics-to-noise ratios, respectively). Our acoustic classification of laughter relied on a somewhat limited set of features and is not comprehensive. Some acoustic dimensions such as sones, jitter, and spectral centroid have only been psychoacoustically calibrated for speech perception, and not for more complex vocalizations such as laughter. Thus, our approach may only capture some of the ways laughter varies acoustically.

Despite variations across laughter studies in methodological choices of social communication categories, speakers, and analytical methods, findings point to consistent ambiguity in volitional laughter with measurable but somewhat minimal acoustic distinctions. A form–function perspective predicts that laughter should manifest itself differently depending on the affective communicative function of the laugh. Our results support this prediction particularly to the extent that volitional laughs similar in valence (e.g., positive) often manifest similar acoustic features and are more likely to be confused with one another. This approach provides an elegant explanation for why we might see both acoustic correlates of specific contexts, yet predictable classification errors.

Previous theorists have referred to this as an affect induction view, referring to the ability of physical sound features in vocalizations to induce an emotional state in listeners, and to signal emotion and motivation as opposed to being functionally referential (Owren and Bachorowski, 2003; Owren and Rendall, 2001; Rendall and Owren, 2010). Our data are clearly aligned with an affect induction approach. It is possible that all categorization accuracy is due to inferences drawn from acoustic patterning that encodes combinations of form–function correspondences. Any accuracy observed in human listeners beyond what is given by such correspondences may be a by-product of the forced-choice paradigm and participants' inferential judgments (i.e., choosing from a list of given categories and making an educated guess). Our results thus support the hypothesis that volitional laughter is a graded rather than discreet signal, not unlike crying in human infants (Bellieni *et al.*, 2004; Gustafson *et al.*, 2000; Koutseff *et al.*, 2018; Lockhart-Bouron *et al.*, 2023; Porter *et al.*, 1986).

In our study, the *malicious* laughter context showed the most distinct acoustic profile and was the most accurately recognized. This may be due to its highly conventionalized and stereotyped features, that may have been shaped by media exposure (Kjeldgaard *et al.*, 2023). Malicious laughter generally reflects a salient negative social intention characterized by high levels of dominance and group exclusion (Kjeldgaard-Christiansen *et al.*, 2023; Nikopoulos, 2017; Ruch *et al.*, 2014). *Mocking* laughter is similar in this way, but aligns more closely with the concept of *schadenfreude*—a joy derived from the misfortunes of others (Nikopoulos, 2017; Szameitat *et al.*, 2009a; Szameitat et al., 2009b; Szameitat et al., 2022). Acoustically, both malicious and mocking laughs are characterized by low fundamental frequency (pitch), loudness, and roughness, sharing features commonly found in vocal intimidation (Anikin *et al.*, 2024).Our findings highlight the complexity of laughter acoustics, wherein acoustic variations may be graded through multiple emotional and motivational dimensions, such as valence and arousal, but also dimensions such as dominance. These dimensional aspects may combine within the same call type in a graded manner, reflecting the nuanced and dynamic nature of human vocal communication. Exploring acoustic gradation across multiple affective dimensions is an important avenue for future research.

In our study, as in most previous work, speakers produced vocalizations in isolation and on command. This experimental protocol is beneficial for many reasons. First, we were interested specifically in *volitional* laughter, that is, laughter that is produced voluntarily. The capacity for such advanced vocal control of nonverbal vocalizations is something that sets humans apart from other primates, and may offer key insights into the social functions of vocal control and modulation (Pisanski *et al.*, 2016). Second, we elicited volitional laughs on demand for experimental control: recordings must be consistent and of sufficient quality for reliable acoustic analysis, and stimuli must not vary in random ways that distract judges in perceptual tasks (e.g., variable in noise, frequency response, etc.). It was also critical to standardize the exact context for each laugh using scripted vignettes, as an "amused" laugh may mean different things to different vocalizers and listeners. However, this control comes with a cost. Volitional vocalizers in the lab generate linguistic and non-linguistic sounds that can differ from spontaneous laughter, including potentially relying on stereotyped preconceptions of communicative acts (e.g., what does sarcasm sound like?) (Bryant and Fox Tree, 2005). The use of professional actors does not particularly alleviate this concern—work has demonstrated that non-professional speakers are often not very different from professional actors (Jürgens *et al.*, 2015). More importantly, real social interaction is generally quite complex, so vocal production occurs in various social and physical contexts that are not cleanly divided in the ways that researchers typically devise their studies. For example, in the case of laughter, a speaker might want to communicate sarcasm, mocking, dominance, and amusement all at the same time. Moreover, in a real-life conversation, a speaker may interject their laugh into a conversational turn with temporal constraints. By removing factors such as these in an experimental context, researchers maximize the likelihood of finding distinctions across gross categories of social interaction, both in production and perception. Thus, while our results indicated that volitional laughter produced in the laboratory can vary slightly as a function of its intended context, volitional laughter in the real world may be less context-specific, with more graded nuances. Moving forward, our research can inform future predictions of how laughter, whether volitional or spontaneous, manifests itself in the often messy world of real social activity.

Future work could also incorporate more elements of spontaneous social interaction into the analytical framework, which can still contain a high degree of volitional laughter. Work using videos from online sources strives to accomplish this but with different costs, such as variations in recording quality, loss of standardization, and lack of information about the vocalizers, the actual contexts in which communicative acts occur, or the extent to which a vocalization was spontaneous or volitional (Anikin and Persson, 2017). Experimental judgment paradigms unfortunately engage cognitive mechanisms that are surely different from those that are triggered during social interaction, with all the dangers of demand characteristics in play (Bryant, 2021). Real-time manipulations of acoustic and visual information (e.g., Arias Sarah *et al.*, 2023), along with subtle methods that probe people's immediate reactions and expectations, could prove to elicit data that better reflect how people navigate their social worlds. We think it is important to explore ways to increase ecological validity in our tasks while maintaining sufficient experimental control.

When people are laughing spontaneously, we can still expect overlap and confusion in laughter production and perception for contexts with a similar valence or function. However, some of the findings in the current study, and other work on volitional laughter, might nevertheless have only limited generalizability for spontaneous laughter given the many specific pragmatic functions that volitional laughs likely serve. Thus, while it is essential to continue studying laughter produced on demand, future work should also examine possible nuances in spontaneous laughter, including how it varies with valence. This approach could deepen our understanding of how volitional laughter may have evolved to exploit the nuances of spontaneous laughter along this valence dimension.

While the laughter contexts used in the current study were informed by past work (Devillers and Vidrascu, 2007; Lavan *et al.*, 2016; Mazzocconi *et al.*, 2020; Nikopoulos, 2017; Ruch *et al.*, 2014; Scott *et al.*, 2014; Szameitat *et al.*, 2022), they are certainly not exhaustive. Our aim was to test the claim that highly differentiated contexts would result in distinct laughter types that would be acoustically and perceptually discriminable, but we tested only eight such contexts. For example, we did not include *backchannel* laughter (i.e., supportive, acknowledging laughter) that is typically very low in arousal. Additionally, some contexts allowed for different interpretations—for instance, in the amused context, the presence of an audience was unclear, which could have influenced laughter. Although we cannot generalize across all possible laughter contexts, we can still largely reject the notion that laughter types are robustly distinct, having instead provided empirical support that laughter operates in a graded way.

A potential limitation in our perception experiment is that vocalizers were French while listeners were English. Pragmatic rules around laughter production (Bryant and Bainbridge, 2022; Mazzocconi *et al.*, 2020) and humor preferences (Martin and Ford, 2018) can vary across cultures, and even closely related societies, but there is no evidence to date that laughter acoustics vary cross-culturally. Moreover, laughter perception seems highly consistent across cultures. For example, participants from disparate cultures can reliably identify spontaneous versus volitional laughs at similar rates (Bryant *et al.*, 2018), universally recognize friends versus strangers colaughing (Bryant *et al.*, 2016), and agree on judgments of emotional dimensions in laughter and other non-linguistic vocalizations (Sauter *et al.*, 2010b). Thus, it is unlikely that the differing cultural backgrounds of our vocalisers and listeners substantially influenced our results, but more research on how laughter may vary across cultures, and potential in-group biases in laughter perception is needed (Kamiloğlu *et al.*, 2022; Szameitat and Szameitat, 2024).

SUPPLEMENTARY MATERIAL

See the supplementary material for supplementary methods and results.

ACKNOWLEDGMENTS

This research is based upon work supported by the French National Research Agency for "SCREAM" Grant No. ANR-21-CE28000701) awarded to D.R. and K.P. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. We thank Mounia Kehy for assisting with data collection, Andrey Anikin for developing *soundgen* and giving feedback on the experimental design and data analysis, and Vincent Arnaud for his assistance with the classification algorithms.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Ethics Approval

Ethical approval for acoustic recording of human subjects and analysis of their non-linguistic vocalizations was provided by the Comité d'Ethique du CHU de Saint-Etienne (IRBN692019/CHUSTE). Informed and written consent was obtained from all participants

DATA AVAILABILITY

Data and codes developed for data processing and analysis are available in zenodo public repository: https://doi.org/10.5281/zenodo.15120255.

Ackermann, H., Hage, S. R., and Ziegler, W. (**2014**). "Brain mechanisms of acoustic communication in humans and nonhuman primates: An evolutionary perspective," Behav. Brain Sci. **37**, 529–546.

Amorim, M. C. P., Anikin, A., Mendes, A. J., Lima, C. F., Kotz, S. A., and Pinheiro, A. P. (**2021**). "Changes in vocal emotion recognition across the life span," Emotion **21**, 315–325.

Anikin, A. (**2019**). "Soundgen: An open-source tool for synthesizing nonverbal vocalizations," Behav. Res. **51**, 778–792.

Anikin, A., Bååth, R., and Persson, T. (**2018**). "Human non-linguistic vocal repertoire: Call types and their meaning," J. Nonverbal Behav. **42**, 53–80.

Anikin, A., and Lima, C. F. (**2017**). "Perceptual and acoustic differences between authentic and acted nonverbal emotional vocalizations," Q. J. Exp. Psychol. **71**, 622–641.

Anikin, A., and Persson, T. (**2017**). "Nonlinguistic vocalizations from online amateur videos for emotion research: A validated corpus," Behav. Res. **49**, 758–771.

Anikin, A., Valente, D., Pisanski, K., Cornec, C., Bryant, G. A., and Reby, D. (**2024**). "The role of loudness in vocal intimidation," J. Exp. Psychol. Gen. **153**, 511–530.

Arias-Sarah, P., Denis, G., Hall, L., Aucouturier, J.-J., Schyns, P., Jack, R. E., and Johansson, P. (**2023**). "DuckSoup: A videoconference experimental platform to transform participants' voice and face in real-time during social interactions," in *2022 Conference on Cognitive Computational Neuroscience*.

Arnaud, V., Pellegrino, F., Keenan, S., St-Gelais, X., Mathevon, N., Levréro, F., and Coupé, C. (**2023**). "Improving the workflow to crack Small, Unbalanced, Noisy, but Genuine (SUNG) datasets in bioacoustics: The case of bonobo calls," PLoS Comput. Biol. **19**, e1010325.

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

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

Bellieni, C. V., Sisto, R., Cordelli, D. M., and Buonocore, G. (**2004**). "Cry features reflect pain intensity in term newborns: An alarm threshold," Pediatr. Res. **55**, 142–146.

Briefer, E. F. (**2012**). "Vocal expression of emotions in mammals: Mechanisms of production and evidence," J. Zool. **288**, 1–20.

Briefer, E. F. (**2020**). "Coding for 'Dynamic' information: Vocal expression of emotional arousal and valence in non-human animals," in *Coding Strategies in Vertebrate Acoustic Communication, Animal Signals and Communication*, edited by T. Aubin and N. Mathison (Springer, Cham, Switzerland), Vol. 7, pp. 137–162.

Bryant, G. (**2020**). "Evolution, structure, and functions of human laughter," in *The Handbook of Communication Science and Biology*, edited by K. Floyd and R. Weber (Routledge, London, UK), pp. 63–77.

Bryant, G. A. (**2021**). "The evolution of human vocal emotion," Emot. Rev. **13**, 25–33.

Bryant, G. A., and Aktipis, C. A. (**2014**). "The animal nature of spontaneous human laughter," Evol. Hum. Behav. **35**, 327–335.

Bryant, G. A., and Bainbridge, C. M. (**2022**). "Laughter and culture," Philos. Trans. R. Soc. B **377**, 20210179.

Bryant, G. A., Fessler D. M., Fusaroli, R., Clint, E., Aarøe, L., Apicella, C. L., Petersen, M. B., Bickham, S. T., Bolyanatz, A., Chavez, B., De Smet, D., Díaz, C., Fančovičová, J., Fux, M., Hu, A., Kamble, S. V., Kameda, T., Li, N. P., Luberti, F. R., Prokop, P., Quintelier, K., Scelza, B. A., Shin, H. J., Soler, M., Stieger, S., Toyokawa, W., van den Hende, E. A., Viciana-Asensio, H., Yildizhan, S. E., Yong, J. C., Yuditha, T., and Zhou, Y. (**2016**). "Detecting affiliation in colaughter across 24 societies," Proc. Natl. Acad. Sci. U.S.A. **113**, 4682–4687.

Bryant, G. A., Fessler, D. M. T., Fusaroli, R., Clint, E., Amir, D., Chávez, B., Denton, K. K., Díaz, C., Duran, L. T., Fanćovićová, J., Fux, M., Ginting, E. F., Hasan, Y., Hu, A., Kamble, S. V., Kameda, T., Kuroda, K., Li, N. P., Luberti, F. R., Díaz, C., Duran, L. T., Fanćovićová, J., Fux, M., Ginting, E. F., Hasan, Y., Hu, A., Kamble, S. V., Kameda, T., Kuroda, K., Li, N. P., Luberti, F. R., Peyravi, R., Prokop, P., Quintelier, K. J. P., Katinka, J. P., Shin, H. J., Stieger, S., Sugiyama, L. S., van den Hende, E. S., Viciana-Asensio, H., Yildizhan, S. E., Yong, J. C., Yuditha, T., and Zhou, Y. (**2018**). "The perception of spontaneous and volitional laughter across 21 societies," Psychol. Sci. **29**, 1515–1525.

Bryant, G. A., and Fox Tree, J. E. (**2005**). "Is there an ironic tone of voice?," Lang. Speech **48**, 257–277.

Bryant G. A. J. (**2023**). "Vocal strategies in verbal irony," in *The Cambridge Handbook of Irony and Thought*, Cambridge Handbooks in Psychology, edited by H. L. Colston and R. W. Gibbs, Jr. (Cambridge University Press, Cambridge, UK), pp. 197–215.

Cartei, V., Banerjee, R., Garnham, A., Oakhill, J., Roberts, L., Anns, S., Bond, R., and Reby, , D. (**2020**). "Physiological and perceptual correlates of masculinity in children's voices," Horm. Behav. **117**, 104616.

Champely, S. (**2020**). "pwr: Basic functions for power analysis," available at https://CRAN.R-project.org/package=pwr.

Charlton, B. D., and Reby, D. (**2016**). "The evolution of acoustic size exaggeration in terrestrial mammals," Nat. Commun. **7**, 12739.

Cosmides, L. (**1983**). "Invariances in the acoustic expression of emotion during speech," J. Exp. Psychol. Hum. Percept. Perform. **9**, 864–881.

Cowen, A. S., Elfenbein, H. A., Laukka, P., and Keltner, D. (**2019**). "Mapping 24 emotions conveyed by brief human vocalization," Am. Psychol. **74**, 698–712.

Curran, W., McKeown, G. J., Rychlowska, M., André, E., Wagner, J., and Lingenfelser, F. (**2018**). "Social context disambiguates the interpretation of laughter," Front. Psychol. **8**, 2342.

Devillers, L., and Vidrascu, L. (**2007**). "Positive and negative emotional states behind the laughs in spontaneous spoken dialogs," in *Interdisciplinary Workshop on The Phonetics of Laughter*, 4-5 August, Saarbrücken, Germany, pp. 37–40.

Engelberg, J. W. M., Schwartz, J. W., and Gouzoules, H. (**2021**). "The emotional canvas of human screams: Patterns and acoustic cues in the perceptual categorization of a basic call type," PeerJ **9**, e10990.

Fernald, A. (**1992**). "Human maternal vocalizations to infants as biologically relevant signals: An evolutionary perspective," in *The Adapted Mind: Evolutionary Psychology and the Generation of Culture* (Oxford University Press, New York), pp. 391–428.

Goeke, C., Finger, H., Diekamp, D., Standvoss, K., and König, P. (**2017**). "LabVanced: A unified JavaScript framework for online studies," in *International Conference on Computational Social Science*, July 10–13, Cologne, Germany.

Goudbeek, M., and Scherer, K. (**2010**). "Beyond arousal: Valence and potency/control cues in the vocal expression of emotion," J. Acoust. Soc. Am. **128**, 1322–1336.

Greenwell, B. M., and Boehmke, B. C. (**2020**). "Variable importance lots–An Introduction to the vip package," R J. **12**, 343.

Gustafson, G., Wood, R., and Green, J. A. (**2000**). "Can we hear the causes of infants' crying?," in *Crying as a Sign, a Symptom, and a Signal* (Cambridge University Press, Cambridge, UK), pp. 8–22.

James, G., Witten, D., Hastie, T., and Tibshirani, R. (**2013**). *An Introduction to Statistical Learning*, Springer Texts in Statistics (Springer Science+Business Media, New York).

Jefferson, G., Sacks, H., and Schegloff, E. (**1987**). "Notes on laughter in the pursuit of intimacy," in *Talk and Social Organisations*, edited by G. Button and J. R. E. Lee (Multilingual Matters, Clevedon, UK), pp. 152–205.

Jürgens, R., Grass, A., Drolet, M., and Fischer, J. (**2015**). "Effect of acting experience on emotion expression and recognition in voice: Non-actors provide better stimuli than expected," J. Nonverbal Behav. **39**, 195–214.

Kamiloğlu, R. G., Fischer, A. H., and Sauter, D. A. (**2020**). "Good vibrations: A review of vocal expressions of positive emotions," Psychon. Bull. Rev. **27**(2), 237–265.

Kamiloğlu, R. G., Tanaka, A., Scott, S. K., and Sauter, D. A. (**2022**). "Perception of group membership from spontaneous and volitional laughter," Philos. Trans. R. Soc. B **377**, 20200404.

Kelly, T., Reby, D., Levréro, F., Keenan, S., Gustafsson, E., Koutseff, A., and Mathevon, N. (**2017**). "Adult human perception of distress in the cries of bonobo, chimpanzee, and human infants," Biol. J. Linn. Soc. **120**, 919–930.

Kjeldgaard-Christiansen, J., Hejná, M., Clasen, M., and Eaton, M. (**2023**). "Evil voices in popular fictions: The case of the exorcist," J. Pop. Cult. **56**, 226–247.

Koutseff, A., Reby, D., Martin, O., Levrero, F., Patural, H., and Mathevon, N. (**2018**). "The acoustic space of pain: Cries as indicators of distress recovering dynamics in pre-verbal infants," Bioacoustics **27**, 313–325.

Lavan, N., Domone, A., Fisher, B., Kenigzstein, N., Scott, S. K., and McGettigan, C. (**2019**). "Speaker sex perception from spontaneous and volitional nonverbal vocalizations," J. Nonverbal Behav. **43**, 1–22.

Lavan, N., Rankin, G., Lorking, N., Scott, S., and McGettigan, C. (**2017**). "Neural correlates of the affective properties of spontaneous and volitional laughter types," Neuropsychologia **95**, 30–39.

Lavan, N., Scott, S. K., and McGettigan, C. (**2016**). "Laugh like you mean it: Authenticity modulates acoustic, physiological and perceptual properties of laughter," J. Nonverbal Behav. **40**, 133–149.

Lima, C. F., Castro, S. L., and Scott, S. K. (**2013**). "When voices get emotional: A corpus of nonverbal vocalizations for research on emotion processing," Behav. Res. **45**, 1234–1245.

Lockhart-Bouron, M., Anikin, A., Pisanski, K., Corvin, S., Cornec, C., Papet, L., Levréro, F., Fauchon, C., Patural, H., Reby, D., and Mathevon N. (**2023**). "Infant cries convey both stable and dynamic information about age and identity," Commun. Psychol. **1**, 26.

Martin, R. A., and Ford, T. (**2018**). *The Psychology of Humor: An Integrative Approach* (Academic Press, Cambridge, MA), 550 pp.

Mazzocconi, C., Tian, Y., and Ginzburg, J. (**2020**). "What's your laughter doing there? A taxonomy of the pragmatic functions of laughter," in IEEE Trans. Affective Comput. **13**, 1302–1321.

Morton E. S.. (**1977**). "On the occurrence and significance of motivation-structural rules in some bird and mammal sounds," Am. Nat. **111**, 855–869.

Nikopoulos, J. (**2017**). "The stability of laughter," Humor **30**, 1–21.

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

Owren, M. J., Amoss, R. T., and Rendall, D. (**2011**). "Two organizing principles of vocal production: Implications for nonhuman and human primates," Am. J. Primatol. **73**, 530–544.

Owren, M. J., and Bachorowski, J. (**2003**). "Reconsidering the evolution of nonlinguistic communication: The case of laughter," J. Nonverbal Behav. **27**, 183–200.

Owren, M. J., and Rendall, D. (**2001**). "Sound on the rebound: Bringing form and function back to the forefront in understanding nonhuman primate vocal signaling," Evol. Anthropol. **10**, 58–71.

Pinheiro, A. P., Anikin, A., Conde, T., Sarzedas, J., Chen, S., Scott, S. K., and Lima, C. F. (**2021**). "Emotional authenticity modulates affective and social trait inferences from voices," Philos. Trans. R. Soc. B **376**, 20200402.

Pisanski, K., and Bryant, G. A. (**2019**). "The evolution of voice perception," in *The Oxford Handbook of Voice Studies*, edited by N. S. Eidsheim and K. Meizel (Oxford University Press, Oxford, UK).

Pisanski, K., Bryant, G. A., Cornec, C., Anikin, A., and Reby, D. (**2022**). "Form follows function in human nonverbal vocalisations," Ethol. Ecol. Evol. **34**, 303–321.

Pisanski, K., Cartei, V., McGettigan, C., Raine, J., and Reby, D. (**2016**). "Voice modulation: A window into the origins of human vocal control?," Trends Cogn. Sci. **20**, 304–318.

Pisanski, K., and Reby, D. (**2021**). "Efficacy in deceptive vocal exaggeration of human body size," Nat. Commun. **12**, 968.

Porter, F. L., Miller, R. H., and Marshall, R. E. (**1986**). "Neonatal pain cries: Effect of circumcision on acoustic features and perceived urgency," Child Dev. **57**, 790–802.

Provine, R. R. (**1993**). "Laughter punctuates speech: Linguistic, social and gender contexts of laughter," Ethology **95**, 291–298.

Provine, R. R. (**2001**). *Laughter: A Scientific Investigation* (Penguin Press, New York).

Raine, J., Pisanski, K., Bond, R., Simner, J., and Reby, D. (**2019a**). "Human roars communicate upper-body strength more effectively than do screams or aggressive and distressed speech," PLoS One **14**, e0213034.

Raine, J., Pisanski, K., Oleszkiewicz, A., Simner, J., and Reby, D. (**2018**). "Human listeners can accurately judge strength and height relative to self from aggressive roars and speech," iScience **4**, 273–280.

Raine, J., Pisanski, K., Simner, J., and Reby, D. (**2019b**). "Vocal communication of simulated pain," Bioacoustics **28**, 404–426.

R Core Team (**2023**). "R: A language and environment for statistical computing [version 4.3.1 (2023-06-16)] Beagle Scouts," (R Foundation for Statistical Computing, Vienna, Austria), https://www.R-project.org.

Rendall, D., and Owren, M. J. (**2010**). "Vocalizations as tools for influencing the affect and behavior of others," in *Handbook of Behavioral Neuroscience*, edited by S. M. Brudzynski (Elsevier, Amsterdam, Netherlands), Vol. 19, pp. 177–185.

Ruch, W. F., Platt, T., Hofmann, J., Niewiadomski, R., Urbain, J., Mancini, M., and Dupont, S. (**2014**). "Gelotophobia and the challenges of implementing laughter into virtual agents interactions," Front. Hum. Neurosci. **8**, 108973.

Rychlowska, M., McKeown, G. J., Sneddon, I., and Curran, W. (**2022**). "The role of contextual information in classifying spontaneous social laughter," J. Nonverbal Behav. **46**(4), 449–466.

Sauter, D. A., Eisner, F., Calder, A. J., and Scott, S. K. (**2010a**). "Perceptual cues in nonverbal vocal expressions of emotion," Q. J. Exp. Psychol. **63**, 2251–2272.

Sauter, D. A., Eisner, F., Ekman, P., and Scott, S. K. (**2010b**). "Cross-cultural recognition of basic emotions through nonverbal emotional vocalizations," Proc. Natl. Acad. Sci. U.S.A. **107**, 2408–2412.

Sauter, D. A., and Scott, S. K. (**2007**). "More than one kind of happiness: Can we recognize vocal expressions of different positive states?," Motiv. Emot. **31**, 192–199.

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

Scott, S. K., Lavan, N., Chen, S., and McGettigan, C. (**2014**). "The social life of laughter," Trends Cogn. Sci. **18**, 618–620.

Sperber, D., and Wilson, D. (**1995**). *Relevance: Communication and Cognition*, 2nd ed. (Blackwell Publishing, Malden, UK).

Szameitat, D. P., Alter, K., Szameitat, A. J., Darwin, C. J., Wildgruber, D., Dietrich, S., and Sterr, A. (**2009a**). "Differentiation of emotions in laughter at the behavioral level," Emotion **9**, 397–405.

Szameitat, D. P., Alter, K., Szameitat, A. J., Wildgruber, D., Sterr, A., and Darwin, C. J. (**2009b**). "Acoustic profiles of distinct emotional expressions in laughter," J. Acoust. Soc. Am. **126**, 354–366.

Szameitat, D. P., and Szameitat, A. J. (**2024**). "Recognition of emotions in German laughter across cultures," Sci. Rep. **14**, 3052.

Szameitat, D. P., Szameitat, A. J., and Wildgruber, D. (**2022**). "Vocal expression of affective states in spontaneous laughter reveals the bright and the dark side of laughter," Sci. Rep. **12**, 5613.

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

Wadewitz, P., Hammerschmidt, K., Battaglia, D., Witt, A., Wolf, F., and Fischer, J. (**2015**). "Characterizing vocal repertoiresôHard vs. soft classification approaches," PLoS One **10**, e0125785.

Wagner, H. L. (**1993**). "On measuring performance in category judgment studies of nonverbal behavior," J. Nonverbal Behav. **17**, 3–28.

Wood, A., Martin, J., and Niedenthal, P. (**2017**). "Towards a social functional account of laughter: Acoustic features convey reward, affiliation, and dominance," PLoS One **12**, e0183811.

Woods, K. J. P., Siegel, M. H., Traer, J., and McDermott, J. H. (**2017**). "Headphone screening to facilitate web-based auditory experiments," Atten. Percept. Psychophys. **79**, 2064–2072.