

Emotional authenticity modulates affective and social trait inferences from voices

Ana P. Pinheiro ^{1*}, Andrey Anikin ^{2,3}, Tatiana Conde ¹, João Sarzedas ¹,
Sinead Chen ⁴, Sophie K. Scott ⁵, César F. Lima ^{5,6*}

¹ CICPSI, Faculdade de Psicologia, Universidade de Lisboa, Lisboa, Portugal

² Equipe de Neuro-Ethologie Sensorielle (ENES) / Centre de Recherche en Neurosciences de Lyon (CRNL),
University of Lyon/Saint-Etienne, CNRS UMR5292, INSERM UMR_S 1028, Saint-Etienne, France

³ Division of Cognitive Science, Lund University, Lund, Sweden

⁴ National Taiwan University, Taipei City, Taiwan

⁵ Institute of Cognitive Neuroscience, University College London, London, UK

⁶ Instituto Universitário de Lisboa (ISCTE-IUL), Lisboa, Portugal

Keywords: voice; emotion; authenticity; social traits; acoustics.

Summary

The human voice is a primary tool for verbal and nonverbal communication. Studies on laughter emphasize a distinction between spontaneous laughter, which reflects a genuinely felt emotion, and volitional laughter, associated with more intentional communicative acts. Listeners can reliably differentiate the two. It remains unclear, however, if they can detect authenticity in other vocalizations, and whether authenticity determines the affective and social impressions that we form about others. Here, 137 participants listened to laughs and cries that could be spontaneous or volitional, and rated them on authenticity, valence, arousal, trustworthiness, and dominance. Bayesian mixed models indicated that listeners detect authenticity similarly well in laughter and crying. Speakers were also perceived to be more trustworthy, and in a higher arousal state, when their laughs and cries were spontaneous. Moreover, spontaneous laughs were evaluated as more positive than volitional ones, and we found that the same acoustic features predicted perceived authenticity and trustworthiness in laughter: high pitch, spectral variability, and less voicing. For crying, associations between acoustic features and ratings were less reliable. These findings indicate that emotional authenticity shapes affective and social trait inferences from voices, and that the ability to detect authenticity in vocalizations is not limited to laughter.

*Authors for correspondence (appinheiro@psicologia.ulisboa.pt; cesar.lima@iscte-iul.pt).

†Present address: Ana P. Pinheiro, Faculdade de Psicologia da Universidade de Lisboa, Alameda da Universidade, 1649-013 Lisboa, Portugal. César F. Lima, Instituto Universitário de Lisboa (ISCTE-IUL), Avenida das Forças Armadas, 1649-026 Lisboa, Portugal.

1. Introduction

The human voice is a rich source of nonverbal information in social interactions. From only a brief clip of someone's voice, we can rapidly extract cues related to aspects such as the age, sex, identity, or emotional state of the speaker. We also form impressions about whether they sound trustworthy or not, or more or less dominant [1]. Most of what we know about voice perception comes from studies using speech stimuli. Examples are the study of emotion perception in speech prosody (e.g., [2–4]) and identity perception in spoken utterances (e.g., [5]). But there is an increasing interest in understanding how we process voices in the absence of concurrent linguistic information, in nonverbal vocalizations such as laughter and crying. Nonverbal vocalizations are distinct from spoken language regarding their underlying articulatory mechanisms [6]. They represent a primitive, universal, and efficient form of emotional communication [7–12].

Recent studies highlight that nonverbal vocalizations can vary a lot, consistent with the complexity and variability that characterize vocal signals [13]. They can vary between speakers, due to differences in the anatomy of the vocal apparatus, for example. They can also vary within the same speaker depending on context. We laugh differently depending on whether we are spontaneously reacting to a funny video, or deliberately trying to show that we agree with our boss in a meeting. A distinction has been made between *spontaneous* and *volitional* vocalizations in several studies on laughter (e.g., [14–17]). Spontaneous laughter is less controlled, reflects a genuinely felt emotion, and is typically reactive to outside events. Volitional laughter is part of more flexible and deliberate communicative acts. It is used to convey appreciation, agreement, or to deceive others [17]. These two forms of laughter might rely on distinct vocal production mechanisms. Spontaneous laughter has been suggested to be initiated by a complex set of midline structures involved in innate vocalizations (e.g., periaqueductal grey [17,18]). Volitional laughter might be supported by the same sensorimotor cortical regions that control the production of learned vocalizations such as speech and song (e.g., lateral motor and premotor cortices [17,18]).

Spontaneous and volitional laughter also differ acoustically and perceptually. Spontaneous laughter is often higher in pitch, longer in duration, and shows spectral features that differ from volitional laughter; volitional laughter, on the other hand, is more nasal than spontaneous laughter [19]. Perceptually, spontaneous laughter is perceived as more authentic than volitional laughter, showing that listeners can distinguish between the two types of vocalizations (e.g., [19–21]). Sensitivity to laughter authenticity has additionally been shown to be consistent across cultures [22], and to relate to distinct cortical responses. McGettigan et al. [23] found that listening to spontaneous laughter elicited increased activation in bilateral superior temporal gyri, whereas listening to volitional laughter elicited increased activation in anterior medial prefrontal and

anterior cingulate cortices, suggesting a more active engagement of mentalizing processes when vocalizations are not genuine.

Two important questions remain unanswered. Because most studies probing the authenticity of nonverbal vocalizations are focused on laughter, it remains unclear whether the reported acoustic and perceptual differences extend to other types of vocalizations. This represents a methodological challenge. It is relatively easy to elicit genuine laughter in the laboratory, but not other vocalizations, namely negative ones such as crying. One study covering a range of positive and negative vocalizations, including crying, found that listeners categorize vocalizations from YouTube videos of emotional episodes as more authentic than acted vocalizations from published corpora [24]. This is suggestive of a general ability to detect the authenticity of vocalizations, although evidence for vocalizations other than laughter remains scarce and conflicting [25]. Furthermore, using stimuli from online videos is not without problems: the quality of the recordings is often low and not comparable to laboratory recordings; the emotions and their authenticity need to be inferred from contextual cues; it is not possible to have the same speakers across emotion and authenticity conditions; and the fact that speakers are being filmed can affect their expressions. We therefore need to determine whether the ability to distinguish genuine and posed vocalizations extends beyond laughter, using well-controlled stimuli that account for these potential confounds.

A second underexplored question is whether the authenticity of vocalizations determines how we form affective and social impressions about a speaker. As for affective evaluations, studies on laughter show that spontaneous vocalizations can lead to higher ratings of valence and arousal [19,23]. Speakers are perceived to be in a more positive and aroused state when they laugh spontaneously compared to when they laugh voluntarily. This needs to be replicated and examined for other vocalizations. As for how authenticity modulates social impressions, nothing is known. Social trait evaluations of faces and voices have been proposed to be based on two core dimensions: trustworthiness and dominance [1,26]. We routinely evaluate whether someone looks or sounds trustworthy or dominant rapidly, often within milliseconds. These judgments are hardly grounded in truth (their accuracy is low), but they have been shown to be relatively consistent across raters, and to affect our decisions, attitudes, and behaviours [27,28]. Studies on facial expressions have shown that people producing Duchenne smiles, which include activation of the muscle that causes wrinkles around the eyes and are associated with genuine happiness, are evaluated as more trustworthy than those producing non-Duchenne smiles [29]. Effects of the Duchenne marker on dominance are less clear [30]. It is plausible that a speaker might be perceived as more trustworthy when their vocalizations are spontaneous compared to when they are volitional.

In the current study, we examined whether listeners detect the authenticity of laughter and crying vocalizations, using well-controlled crying stimuli generated via emotion induction in a laboratory setting. Based on previous findings [19–21,24], we predicted that spontaneous vocalizations would be associated with higher authenticity ratings. We also asked whether emotional authenticity shapes how listeners evaluate the affective state of the speaker, namely valence and arousal, as well as their trustworthiness and dominance. Spontaneous vocalizations were expected to be rated higher in arousal and more extreme in valence: more positive in the case of spontaneous laughter, and more negative in the case of spontaneous crying. Based on findings from smile authenticity [29], spontaneous vocalizations were also expected to produce higher perceptions of trustworthiness. As for potential effects on impressions of dominance, our approach was exploratory. Finally, we examined acoustic differences between spontaneous and volitional vocalizations, and how acoustic features predicted subjective ratings.

2. Methods

2.1. Participants

One hundred thirty-seven volunteers participated in the study ($M_{\text{age}} = 21.64$ years, $SD = 6.13$, range = 19-57 years; 115 female). All were European Portuguese native speakers, had normal hearing, and normal or corrected-to-normal visual acuity.

The study was approved by the ethics committee of Faculty of Psychology – University of Lisbon. Before taking part, all participants were informed about the procedures and provided written informed consent. They received course credit for their participation.

We used Bayesian inference in our analyses, which relies on estimates of uncertainty and not on p values. Nevertheless, our sample size can also be considered appropriate according to the standard null hypothesis significance testing approach. An *a priori* power analysis with G*Power 3.1 [31] indicated that a sample size of at least 84 would be required to detect significant correlations of $r = .30$ or larger between variables, considering an alpha level of .05 and a power of .80.

2.2. Stimuli

The experimental stimuli consisted of 75 vocalizations, divided into four conditions: 19 spontaneous laughs, 19 volitional laughs, 19 spontaneous cries, and 18 volitional cries. They were selected from a larger set of stimuli recorded by six speakers (three women) within a sound-proof anechoic chamber at University College London. Laughter vocalizations have been used in previous behavioural and neuroimaging experiments focused on authenticity detection [19–21,32]. Crying vocalizations have also been used in prior studies [21,33], but this is the first one to address how they are perceived regarding their authenticity.

To record volitional laughter and crying, the six speakers were asked to intentionally produce these vocal expressions in the absence of a corresponding emotional eliciting event, and to make them sound as natural and credible as possible. This is in line with the procedure typically used for the recording of acted stimuli [34–37].

As for genuine vocalizations, spontaneous laughter was elicited using an amusement induction procedure in a social interactive setting: speakers watched video clips, which they had previously identified as amusing and which would easily make them laugh. The experimenters knew the speakers well and interacted with them during the recording session to promote the naturalness and the social nature of the laughs. Spontaneous crying was also obtained via an emotion induction procedure. Speakers were asked to recall difficult (upsetting) past episodes and/or to initially produce volitional crying to promote a transition into spontaneous crying reflecting genuine experience of sadness. Feelings of amusement and sadness throughout and after recording the corresponding spontaneous expressions were reported by all speakers.

2.3. Procedure

The experiment was conducted in a quiet room in a laboratory setting at the Faculty of Psychology, University of Lisbon. Participants were tested in small group sessions with up to 8 participants per session.

Vocal stimuli were presented via headphones, and stimulus presentation and response recording were controlled using *Qualtrics* software. Participants were instructed to rate the sounds as quickly as possible, following their first impressions. After the presentation of each vocalization, participants were asked to indicate the emotion that better characterized the sound in a three-alternative forced-choice categorization: ‘sadness’, ‘neutral’, or ‘happiness’. They then rated the vocalization regarding the dimensions of emotional authenticity, valence, and arousal, as well as the social traits of trustworthiness and dominance. Nine-point scales were used, from 1 (minimum) to 9 (maximum).

Vocalizations were presented once, in a pseudo-random order to avoid the presentation of more than two consecutive vocalizations from the same category. Before the experiment, two practice examples were provided: a crying and a laughter exemplar from the Montreal Affective Voices [34]. The session lasted around 45 minutes.

2.4. Data Analysis

2.4.1. Behavioural data

Statistical analyses were performed on unaggregated responses from individual trials using Bayesian mixed models and the *brms* R package [38]. All results were summarized as the medians of posterior distributions and 95% credible intervals (CI). When contrasting two conditions (e.g.,

spontaneous vs. volitional vocalizations), the CI includes the most credible values for the difference given the data and the model and, if that does not include 0, we can infer that there is evidence in favour of an actual difference between conditions (see Supplement). The code used for data analysis and the full data set can be found here: https://osf.io/57syv/?view_only=c98e91f70a2d49e8ad902bbde2a4482c

2.4.2. Acoustic features

The audio files were downsampled to 22050 Hz, high-passed over 90 Hz to remove low-frequency noise, and analysed acoustically with the *soundgen* R package [39]. Intonation contours were manually verified and, if necessary, corrected using the *pitch_app()* interactive environment. When appropriate, acoustic descriptives were summarized as the median value and standard deviation across the entire sound duration. Among the potentially large number of quantifiable acoustic characteristics, we focused on nine key variables, chosen a priori because they are theoretically meaningful, have often been reported in earlier studies, and can be measured reliably (e.g., [24]):

- a) Duration (s): the duration of a stimulus without counting silent frames at the beginning and end;
- b) Harmonics-to-noise ratio (HNR), median and *SD* (dB): the harmonics-to-noise ratio, a measure of pitch quality or tonality calculated only for voiced frames;
- c) Novelty: a measure of spectral variability derived from the self-similarity matrix (SSM) of a vocalization by sliding a 200-ms Gaussian checkerboard matrix along the SSM's diagonal;
- d) Pitch, median and *SD* (Hz): manually verified fundamental frequency or perceived tone height;
- e) Spectral centroid, median and *SD* (Hz): the first spectral moment or center of gravity of the spectrum of voiced frames, which perceptually corresponds to timbral brightness;
- f) Voiced (%): the proportion of voiced frames.

The variables measured in Hz were transformed to the more perceptually relevant logarithmic scale, following which all variables were scaled to have a mean of 0 and *SD* of 1. We used median rather than mean values, since medians are more robust to outliers, such as frames with incorrectly measured pitch or external noise.

To test the effect of acoustic predictors on the ratings, we again used multivariate ordinal regression and predicted the ratings on all five scales as a function of the nine measured acoustic features (see Supplement).

3. Results

3.1. *Affective and social ratings of spontaneous and volitional vocalizations*

As a measure of inter-rater agreement, we aggregated the ratings of each sound on each of the five scales and calculated the mean Pearson's correlation between the responses of each participant and these aggregated ratings. Within each vocalization type (laughter and crying), correlations ranged from .44 to .62 for four scales: authenticity, valence, arousal, and trustworthiness. They were considerably lower for the dominance scale (.37 for laughter and .20 for crying). Likewise, the intraclass correlation coefficient (ICC), estimated using a two-way random model and absolute agreement, revealed lower reliability for the dominance scale ($< .1$) compared to the other four scales (.1-.3). Participants thus found it more difficult to rate the stimuli on dominance and, therefore, the results for this scale should be treated with caution.

The accuracy of emotion recognition in the forced-choice classification task was above 95% (see Supplement). Figure 1 shows how spontaneous and volitional laughter and crying were rated on all five scales. In line with our prediction, spontaneous vocalizations were rated as 1.72 points more authentic than volitional ones (95% CI [1.35, 2.09]). The difference was 2.03 points for laughter (95% CI [1.52, 2.54]) and 1.40 points for crying (95% CI [0.87, 1.94]). The difference between laughter and crying in terms of this authenticity contrast was not statistically robust (0.63 points higher for laughter, 95% CI [-0.13, 1.36]). Laughter was overall judged to be slightly more authentic than crying (0.66 points, 95% CI [0.29, 1.03]).

If we consider that an authenticity rating is 'correct' when it is above the midpoint of the 1-9 scale (>5), and below in the case of volitional vocalizations, the overall accuracy of recognizing authenticity was 60.6% (95% CI [55.6, 65.7]). This varied across vocalization types: 65.8% for spontaneous laughter (95% CI [55.2, 75.6]), 62.3% for volitional laughter (95% CI [51.3, 71.8]), 41.9% for spontaneous crying (95% CI [31.6, 52.9]), and 72.6% for volitional crying (95% CI [62.7, 81.4]). Despite the slight bias to treat laughter as more authentic than crying, there were no statistically robust differences in accuracy of authenticity detection when comparing volitional with spontaneous stimuli (13.4% higher accuracy for volitional, 95% CI [-2.8, 29.6]), or laughs with cries (6.7% higher accuracy for laughs, 95% CI [-9.5, 22.8]).

Regarding valence ratings, laughter was generally rated as more positive (6.10, 95% CI [5.88, 6.32]) than crying (3.16, 95% CI [2.96, 3.37]), as expected (see Figure 1). Authenticity also played a role in the case of laughter. Spontaneous laughs were rated as 0.93 points more positive than volitional laughs (95% CI [0.53, 1.34]). No such effect was found for crying, for which the difference was only 0.03 points (95% CI [-0.36, 0.4]). In general, spontaneous vocalizations were also rated as 1.04 points (95% CI [0.8, 1.29]) more arousing compared to volitional vocalizations, by 1.55 points for laughs (95% CI [1.21, 1.91]), and 0.53 for cries (95% CI [0.18, 0.9]). Averaging

across spontaneous and volitional stimuli, laughter was rated as 0.38 points (95% CI [0.13, 0.62]) higher on arousal compared to crying.

As for trait inferences, spontaneous expressions were perceived as more trustworthy. The difference in the ratings was 1.39 points for laughter (95% CI [1.02, 1.76]), and 0.97 for crying (95% CI [0.58, 1.36]; see Figure 1, right panel). The effect of authenticity on trustworthiness ratings was similar for laughter and crying (0.42 higher for laughter, 95% CI [-0.13, 0.95]). Averaging across spontaneous and volitional vocalizations, trustworthiness ratings were 0.48 points higher for laughter compared to crying (95% CI [0.22, 0.75]). No differences were found for dominance ratings: inferences were similar for spontaneous and volitional vocalizations (a difference of 0.15 points for laughter, 95% CI [-0.09, 0.39], and of 0.02 points for crying, 95% CI [-0.23, 0.27]).

In sum, spontaneous vocalizations were rated as more authentic, arousing, and trustworthy than volitional ones. Spontaneous laughs were rated as more positive than their volitional counterparts. Emotional authenticity did not affect valence perception for crying, and it did not affect inferences of dominance for both laughter and crying.

<INSERT FIGURE 1 HERE>

3.2. Predicting social inferences from affective ratings

Correlations among the five scales were small-to-moderate (all r s < .3 for crying and < .65 for laughter). However, there was a strong linear relationship between authenticity and trustworthiness ratings, $r = .83$. Thus, both ‘objective’ (stimulus-based) and perceived authenticity were associated with inferences of higher trustworthiness. Notably, correlations between authenticity and arousal, a variable that has been highlighted as a potential marker of authenticity [16,19], were much lower both for laughter ($r = .42$) and for crying ($r = .27$).

To model the relationship between the five scales, we predicted social ratings (trust and dominance) from authenticity, valence, and arousal ratings. Authenticity strongly predicted trustworthiness ratings in laughter (5.10 points higher trustworthiness for an increase in authenticity ratings from 1 to 9, 95% CI [4.78, 5.42]) and crying (5.41, 95% CI [5.09, 5.71]; see Figure 2). Valence ratings also predicted trustworthiness, but the effect was much smaller, and limited to laughter (1.51 points, 95% CI [1.27, 1.77]; for crying, 0.14 points, 95% CI [-0.06, 0.34]). Arousal ratings were not credibly related to trustworthiness in either laughter (-0.11 points, 95% CI [-0.3, 0.08]) or crying (0.05 points, 95% CI [-0.12, 0.22]).

Perceived authenticity in laughter predicted evaluations of dominance, albeit weakly (0.6 points, 95% CI [0.26, 0.93]), but not in crying (-0.21 points, 95% CI [-0.62, 0.17]). Perceived arousal and valence in both laughter and crying also predicted dominance (effect sizes ~1 point on a 1-9 scale, see Figure 2).

<INSERT FIGURE 2 HERE>

3.3. Predicting affective and social ratings from acoustic features of vocalizations

As described in Methods, we focused on nine theoretically meaningful acoustic characteristics of nonverbal vocalizations. Differences were found between spontaneous and volitional vocalizations: spontaneous laughs had higher pitch, harmonics-to-noise ratio, brighter timbre, and greater general variability; spontaneous cries had higher pitch, variable timbral brightness, and less voicing (see Supplement). We tested how acoustic features affected the ratings on each of five scales, separately for laughter and crying (Figure 3). Because many of the acoustic predictors are correlated, we estimated their partial effects in multiple regressions. High-pitched and generally variable laughs were judged to be more authentic and trustworthy. This was indicated by the positive effects of median pitch and novelty on authenticity and trustworthiness ratings: 0.62 (95% CI [0.16, 1.08]) and 0.5 (95% CI [0.17, 0.8]) increase in authenticity and trustworthiness ratings, respectively, for a 1-SD increase in pitch; and 0.76 (95% CI [0.28, 1.18]) and 0.56 (95% CI [0.24, 0.85]) increase in authenticity and trustworthiness, respectively, for a 1-SD increase in novelty. Laughs with a smaller proportion of voiced frames were also judged to be more authentic (0.56, 95% CI [0.13, 0.96]).

A shift of spectral energy towards higher harmonics, indicative of increased vocal effort and a bright voice (spectral centroid), predicted higher perceived arousal (0.55, 95% CI [0.25, 0.82]). It also predicted higher authenticity, although the effect did not clear the ROPE (0.50, 95% CI [0.09, 0.89]). More positive valence in laughs was primarily predicted by greater novelty (0.46, 95% CI [0.2, 0.7]), although there were also statistically uncertain effects of pitch (0.31, 95% CI [0.05, 0.56]) and spectral centroid (0.25, 95% CI [0.03, 0.47]).

In sum, laughs were perceived to be more authentic when they were high-pitched, very variable, and had a lot of unvoiced fragments. Increased vocal effort signalled high arousal and to some extent authenticity as well. We did not observe any statistically robust acoustic predictors of dominance ratings in laughs, or of any ratings in cries.

<INSERT FIGURE 3 HERE>

4. Discussion

Our findings confirm that the acoustic and perceptual differences between spontaneous and volitional vocalizations extend to vocalizations of negative valence. We observed that the generation of laughs and cries in spontaneous and volitional contexts was effective in leading to changes in perceived authenticity. That is, listeners could reliably judge whether a laugh or a cry was produced in response to genuine amusement or sadness, respectively, or whether they were

produced in the absence of an emotional state. This capacity to classify spontaneous and volitional vocal expressions confers advantages in social interactions, for example to avoid deception [40].

Authenticity shaped the perception of the affective qualities of vocalizations, in good agreement with previous studies on laughter (e.g., [14,19,23]). Specifically, spontaneous laughs (but not cries) were perceived as more pleasant than their volitional counterparts. Both spontaneous laughs and cries were perceived as more arousing (see also [19]). We also showed that authenticity affects how listeners form social impressions about a speaker. Social trait inference has been found to interact with emotional perception [1,41–43]. For example, we often rely on transient signals (e.g., the emotional quality of the voice) to make inferences of more stable characteristics of a speaker, such as whether they are trustworthy or friendly (over-generalization hypothesis [44,45]). We extend these findings by showing that authenticity strongly predicts how trustworthy a voice is perceived to be, irrespective of vocalization type.

Of note, the weight of affective cues in social trait evaluation differed for trustworthiness and dominance. Trustworthiness was strongly predicted by authenticity, less by valence (and only in laughter), and *not* by arousal. These findings document for the first time that authenticity might be signalling something unique, that arousal alone does not. Authenticity often correlates with arousal [16,19], as increased arousal has been linked to the presence of “hard-to-fake” properties of voices, but our findings suggest they reflect dissociable dimensions of vocalizations. Supporting this notion, studies with facial expressions indicate that spontaneous and acted smiles are accurately discriminated even when the stimuli are matched for arousal [46]. The links between pleasantness (valence) and trustworthiness of vocalizations confirms the alignment of the two variables in the two-dimensional social voice space [1,47], as they are both related to approachability in social interactions. In contrast, prediction of social inference from affective ratings was generally less robust in the case of dominance, which could be related to the lower agreement among participants when judging how dominant a speaker sounded like. Alternatively, these findings may highlight the primary role of trustworthiness (relative to dominance) in social evaluations [48,49].

The current study also identified which acoustic characteristics predict affective and social judgements of speakers. Authenticity in laughs was predicted by more high-frequency energy (i.e., higher pitch and spectral slope) and variability (novelty). There was a considerable overlap in the acoustic features predicting authenticity and trustworthiness ratings in laughs, consistent with the strong correlation between the two rating scales. This suggests that participants might be partly using the same ‘code’ to make inferences about these two aspects of laughter. The link between authenticity and trustworthiness might also be related to the social function of laughter, namely in establishing and maintaining social bonds [17,50,51], which may occur via increased trust in others. In previous studies, higher pitch has been shown to strongly affect social inference and,

specifically, to be associated with higher trustworthiness (e.g., [52]), namely mitigating the aversiveness of speech antisocial cues [53]. Higher pitch and more variable acoustic parameters have also been previously associated with trustworthiness/valence [1,52] and with increased trusting behaviours toward the speaker [54].

Affective and social ratings of cries could not be related to the acoustic features tested here. This suggests that the acoustic hallmarks of authenticity partly differ for positive (laughs) and negative (cries) nonverbal vocalizations. Since our analysis only focused on nine acoustic features, one possibility is that these specific cues play a more subtle role when predicting affective and social evaluations of cries. This could also explain the smaller difference in perceived authenticity between spontaneous and volitional crying compared to laughter. Studies including more stimuli and testing other acoustic parameters are therefore warranted.

5. Conclusions

The present study provides evidence that listeners can reliably infer the emotional authenticity of laughter and crying sounds. It also indicates that emotional authenticity shapes how listeners evaluate the affective state of a speaker, in terms of valence and arousal, as well as how they make social trait inferences, namely regarding trustworthiness. We provide the first demonstration that spontaneous vocal expressions are perceived to be more trustworthy. Moreover, we show that spontaneous vocal expressions differ from volitional ones in several acoustic features, and that the constellation of acoustic differences is partly unique for laughter and crying. It was difficult to predict the ratings of cries from acoustic features, but for laughter we could see that the acoustic predictors of perceived authenticity and trustworthiness are similar. High-pitched and acoustically variable laughs, with a lot of unvoiced fragments, were considered to be both more genuine and more trustworthy.

Our findings have implications for theories of social perception (e.g., [55]). They indicate that authenticity should be considered when accounting for how listeners form social impressions from voices. They raise the interesting possibility that genuine vocal expressions may lead to more trusting, cooperative and prosocial behaviour in social interactions, a hypothesis that needs to be addressed in future studies.

Acknowledgments

This work was supported by Fundação para a Ciência e a Tecnologia, Portugal (FCT; grant number PTDC/MHC-PCN/0101/2014 awarded to APP).

References

1. McAleer P, Todorov A, Belin P. 2014 How do you say ‘hello’? Personality impressions from brief novel voices. *PLoS One* **9**, e90779. (doi:10.1371/journal.pone.0090779)
2. Schirmer A, Kotz SA. 2006 Beyond the right hemisphere: Brain mechanisms mediating vocal emotional processing. *Trends Cogn. Sci.* **10**, 24–30. (doi:10.1016/j.tics.2005.11.009)
3. Grandjean D. 2020 Brain Networks of Emotional Prosody Processing. *Emot. Rev.* (doi:10.1177/1754073919898522)
4. Pinheiro AP, Del Re E, Mezin J, Nestor PG, Rauber A, McCarley RW, Gonçalves OF, Niznikiewicz MA. 2013 Sensory-based and higher-order operations contribute to abnormal emotional prosody processing in schizophrenia: An electrophysiological investigation. *Psychol. Med.* **43**, 603–618. (doi:10.1017/S003329171200133X)
5. Lavan N, Burston LFK, Garrido L. 2019 How many voices did you hear? Natural variability disrupts identity perception from unfamiliar voices. *Br. J. Psychol.* **110**, 576–593. (doi:10.1111/bjop.12348)
6. Scott SK, Sauter D, McGettigan C. 2010 Brain mechanisms for processing perceived emotional vocalizations in humans. In *Handbook of Behavioral Neuroscience* (ed SM Brudzynski), pp. 187–197. (doi:10.1016/B978-0-12-374593-4.00019-X)
7. Castiajo P, Pinheiro AP. 2019 Decoding emotions from nonverbal vocalizations: How much voice signal is enough? *Motiv. Emot.* **43**, 803–813. (doi:10.1007/s11031-019-09783-9)
8. Lima CF, Anikin A, Monteiro AC, Scott SK, Castro SL. 2018 Automaticity in the recognition of nonverbal emotional vocalizations. *Emotion* **19**, 219–233. (doi:10.1037/emo0000429)
9. Pinheiro AP, Lima D, Albuquerque PB, Anikin A, Lima CF. 2019 Spatial location and emotion modulate voice perception. *Cogn. Emot.* **33**. (doi:10.1080/02699931.2019.1586647)
10. Sauter DA, Eisner F, Calder AJ, Scott SK. 2010 Perceptual cues in nonverbal vocal expressions of emotion. *Q. J. Exp. Psychol.* **63**, 2251–2272. (doi:10.1080/17470211003721642)
11. Sauter DA, Eisner F, Ekman P, Scott SK. 2010 Cross-cultural recognition of basic emotions through nonverbal emotional vocalizations. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 2408–2412. (doi:10.1073/pnas.0908239106)
12. Sauter DA, Crasborn O, Engels T, Kamiloglu RG, Sun R, Eisner F, Haun DBM. 2020 Human Emotional Vocalizations Can Develop in the Absence of Auditory Learning. *Emotion* **20**, 1435–1445. (doi:10.1037/emo0000654)

13. Lavan N, Burton AM, Scott SK, McGettigan C. 2019 Flexible voices: Identity perception from variable vocal signals. *Psychon. Bull. Rev.* **26**, 90–102. (doi:10.3758/s13423-018-1497-7)
14. Bryant GA, Aktipis CA. 2014 The animal nature of spontaneous human laughter. *Evol. Hum. Behav.* **35**, 327–335. (doi:10.1016/j.evolhumbehav.2014.03.003)
15. Gervais M, Wilson DS. 2005 The evolution and functions of laughter and humor: A synthetic approach. *Q. Rev. Biol.* **80**, 395–430. (doi:10.1086/498281)
16. McKeown G, Sneddon I, Curran W. 2015 Gender differences in the perceptions of genuine and simulated laughter and amused facial expressions. *Emot. Rev.* **7**, 30–38. (doi:10.1177/1754073914544475)
17. Scott SK, Lavan N, Chen S, McGettigan C. 2014 The social life of laughter. *Trends Cogn. Sci.* **18**, 618–620. (doi:10.1016/j.tics.2014.09.002)
18. Pisanski K, Cartei V, McGettigan C, Raine J, Reby D. 2016 Voice Modulation: A Window into the Origins of Human Vocal Control? *Trends Cogn. Sci.* **20**, 304–318. (doi:10.1016/j.tics.2016.01.002)
19. Lavan N, Scott SK, McGettigan C. 2016 Laugh Like You Mean It: Authenticity Modulates Acoustic, Physiological and Perceptual Properties of Laughter. *J. Nonverbal Behav.* **40**, 133–149. (doi:10.1007/s10919-015-0222-8)
20. Neves L, Cordeiro C, Scott SK, Castro SL, Lima CF. 2018 High emotional contagion and empathy are associated with enhanced detection of emotional authenticity in laughter. *Q. J. Exp. Psychol.* **71**, 2355–2363. (doi:10.1177/1747021817741800)
21. O’Nions E, Lima CF, Scott SK, Roberts R, McCrory EJ, Viding E. 2017 Reduced Laughter Contagion in Boys at Risk for Psychopathy. *Curr. Biol.* **27**, 3049–3055. (doi:10.1016/j.cub.2017.08.062)
22. Bryant GA *et al.* 2018 The Perception of Spontaneous and Volitional Laughter Across 21 Societies. *Psychol. Sci.* **29**, 1515–1525. (doi:10.1177/0956797618778235)
23. McGettigan C, Walsh E, Jessop R, Agnew ZK, Sauter DA, Warren JE, Scott SK. 2015 Individual differences in laughter perception reveal roles for mentalizing and sensorimotor systems in the evaluation of emotional authenticity. *Cereb. Cortex* **25**, 246–257. (doi:10.1093/cercor/bht227)
24. Anikin A, Lima CF. 2018 Perceptual and acoustic differences between authentic and acted nonverbal emotional vocalizations. *Q. J. Exp. Psychol.* **71**, 622–641. (doi:10.1080/17470218.2016.1270976)
25. Atias D, Aviezer H. 2020 Real-life and posed vocalizations to lottery wins differ fundamentally in their perceived valence. *Emotion* (doi:10.1037/emo0000931)
26. Oosterhof NN, Todorov A. 2008 The functional basis of face evaluation. *Proc. Natl.*

- Acad. Sci. U. S. A.* **105**, 11087–11092. (doi:10.1073/pnas.0805664105)
27. Todorov A, Olivola CY, Dotsch R, Mende-Siedlecki P. 2015 Social attributions from faces: Determinants, consequences, accuracy, and functional significance. *Annu. Rev. Psychol.* **66**, 519–545. (doi:10.1146/annurev-psych-113011-143831)
 28. Todorov A, Said CP, Engell AD, Oosterhof NN. 2008 Understanding evaluation of faces on social dimensions. *Trends Cogn. Sci.* **12**, 455–460. (doi:10.1016/j.tics.2008.10.001)
 29. Gunnery SD, Ruben MA. 2016 Perceptions of Duchenne and non-Duchenne smiles: A meta-analysis. *Cogn. Emot.* **30**, 501–515. (doi:10.1080/02699931.2015.1018817)
 30. Quadflieg S, Vermeulen N, Rossion B. 2013 Differential Reliance on the Duchenne Marker During Smile Evaluations and Person Judgments. *J. Nonverbal Behav.* **37**, 69–77. (doi:10.1007/s10919-013-0147-z)
 31. Faul F, Erdfelder E, Lang A-G, Buchner A. 2007 G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods* **39**, 175–191. (doi:10.3758/BF03193146)
 32. Lima CF, Brancatisano O, Fancourt A, Müllensiefen D, Scott SK, Warren JD, Stewart L. 2016 Impaired socio-emotional processing in a developmental music disorder. *Sci. Rep.* **6**, 34911. (doi:10.1038/srep34911)
 33. Lavan N, Lima CF, Harvey H, Scott SK, McGettigan C. 2015 I thought that I heard you laughing: Contextual facial expressions modulate the perception of authentic laughter and crying. *Cogn. Emot.* **29**, 935–944. (doi:10.1080/02699931.2014.957656)
 34. Belin P, Fillion-Bilodeau S, Gosselin F. 2008 The Montreal Affective Voices: A validated set of nonverbal affect bursts for research on auditory affective processing. *Behav. Res. Methods* **40**, 531–539. (doi:10.3758/BRM.40.2.531)
 35. Maurage P, Joassin F, Philippot P, Campanella S. 2007 A validated battery of vocal emotional expressions. *Neuropsychol. Trends* **2**, 63–74.
 36. Lima CF, Castro SL, Scott SK. 2013 When voices get emotional: A corpus of nonverbal vocalizations for research on emotion processing. *Behav. Res. Methods* **45**, 1234–1245. (doi:10.3758/s13428-013-0324-3)
 37. Amorim M, Anikin A, Mendes AJ, Lima CF, Kotz SA, Pinheiro AP. 2019 Changes in Vocal Emotion Recognition Across the Life Span. *Emotion* (doi:10.1037/emo0000692)
 38. Bürkner PC. 2018 Advanced Bayesian multilevel modeling with the R package brms. *R J.*, arXiv preprint arXiv:1705.11123. (doi:10.32614/rj-2018-017)
 39. Anikin A. 2019 Soundgen: An open-source tool for synthesizing nonverbal vocalizations. *Behav. Res. Methods* **51**, 778–792. (doi:10.3758/s13428-018-1095-7)
 40. Lavan N, Rankin G, Lorking N, Scott S, McGettigan C. 2017 Neural correlates of the affective properties of spontaneous and volitional laughter types. *Neuropsychologia* **95**,

- 30–39. (doi:10.1016/j.neuropsychologia.2016.12.012)
41. Berry DS. 1990 Vocal attractiveness and vocal babyishness: Effects on stranger, self, and friend impressions. *J. Nonverbal Behav.* **14**. (doi:10.1007/BF00996223)
 42. Hughes SM, Dispenza F, Gallup GG. 2004 Ratings of voice attractiveness predict sexual behavior and body configuration. *Evol. Hum. Behav.* **25**, 295–304. (doi:10.1016/j.evolhumbehav.2004.06.001)
 43. Montepare JM, Zebrowitz-McArthur L. 1987 Perceptions of adults with childlike voices in two cultures. *J. Exp. Soc. Psychol.* **23**, 331–349. (doi:10.1016/0022-1031(87)90045-X)
 44. McArthur LZ, Baron RM. 1983 Toward an ecological theory of social perception. *Psychol. Rev.* **90**, 215–238. (doi:10.1037//0033-295x.90.3.215)
 45. Zebrowitz LA, Collins MA. 1997 Accurate social perception at zero acquaintance: The affordances of a Gibsonian approach. *Personal. Soc. Psychol. Rev.* **1**, 204–223. (doi:10.1207/s15327957pspr0103_2)
 46. Murphy NA, Lehrfeld JM, Isaacowitz DM. 2010 Recognition of Posed and Spontaneous Dynamic Smiles in Young and Older Adults. *Psychol. Aging* **25**, 811–821. (doi:10.1037/a0019888)
 47. Oosterhof NN, Todorov A. 2009 Shared Perceptual Basis of Emotional Expressions and Trustworthiness Impressions From Faces. *Emotion* **9**, 128–133. (doi:10.1037/a0014520)
 48. Cuddy AJC, Fiske ST, Glick P. 2008 Warmth and Competence as Universal Dimensions of Social Perception: The Stereotype Content Model and the BIAS Map. *Adv. Exp. Soc. Psychol.* **40**, 61–149. (doi:10.1016/S0065-2601(07)00002-0)
 49. Sutherland CAM, Rhodes G, Burton NS, Young AW. 2020 Do facial first impressions reflect a shared social reality? *Br. J. Psychol.* **111**, 215–232. (doi:10.1111/bjop.12390)
 50. Warren JE, Sauter DA, Eisner F, Wiland J, Dresner MA, Wise RJS, Rosen S, Scott SK. 2006 Positive emotions preferentially engage an auditory-motor ‘mirror’ system. *J. Neurosci.* **26**, 13067–13075. (doi:10.1523/JNEUROSCI.3907-06.2006)
 51. Provine RR. 2004 Laughing, tickling, and the evolution of speech and self. *Curr. Dir. Psychol. Sci.* **13**, 215–218. (doi:10.1111/j.0963-7214.2004.00311.x)
 52. Ponsot E, Burred JJ, Belin P, Aucouturier JJ. 2018 Cracking the social code of speech prosody using reverse correlation. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 3972–3977. (doi:10.1073/pnas.1716090115)
 53. O’Connor JJM, Barclay P. 2018 High voice pitch mitigates the aversiveness of antisocial cues in men’s speech. *Br. J. Psychol.* **109**, 812–829. (doi:10.1111/bjop.12310)
 54. Torre I, White L, Goslin J. 2016 Behavioural mediation of prosodic cues to implicit judgements of trustworthiness. In *Proceedings of the International Conference on Speech Prosody*, pp. 816–820. Baixas. (doi:10.21437/speechprosody.2016-167)

55. Young AW, Frühholz S, Schweinberger SR. 2020 Face and Voice Perception: Understanding Commonalities and Differences. *Trends Cogn. Sci.* **24**, 398–410. (doi:10.1016/j.tics.2020.02.001)

Figure captions

Figure 1. Ratings on each of five scales as a function of the spontaneous or volitional nature of the rated vocalizations (A) and the difference in ratings between spontaneous and volitional vocalizations (B). The points and error bars show fitted values (medians of posterior distribution with 95% CI), and violin plots show the distribution of mean observed values per stimulus.

Figure 2. Effects of authenticity, valence, and arousal ratings on the perceived trustworthiness and dominance of the speaker, on a 1-9 scale in multiple regression. Medians of posterior distributions and 95% CIs.

Figure 3. Predicting ratings on the five scales from acoustic characteristics: medians of posterior distribution and 95% CIs from multivariate ordinal regression. The highlighted effects clear the Region of Practical Equivalence (ROPE) of $(-0.1, 0.1)$, corresponding to an effect of at least 0.1 on the rating scale (1 to 9) when changing the predictor by 1 SD. HNR = harmonics-to-noise ratio.

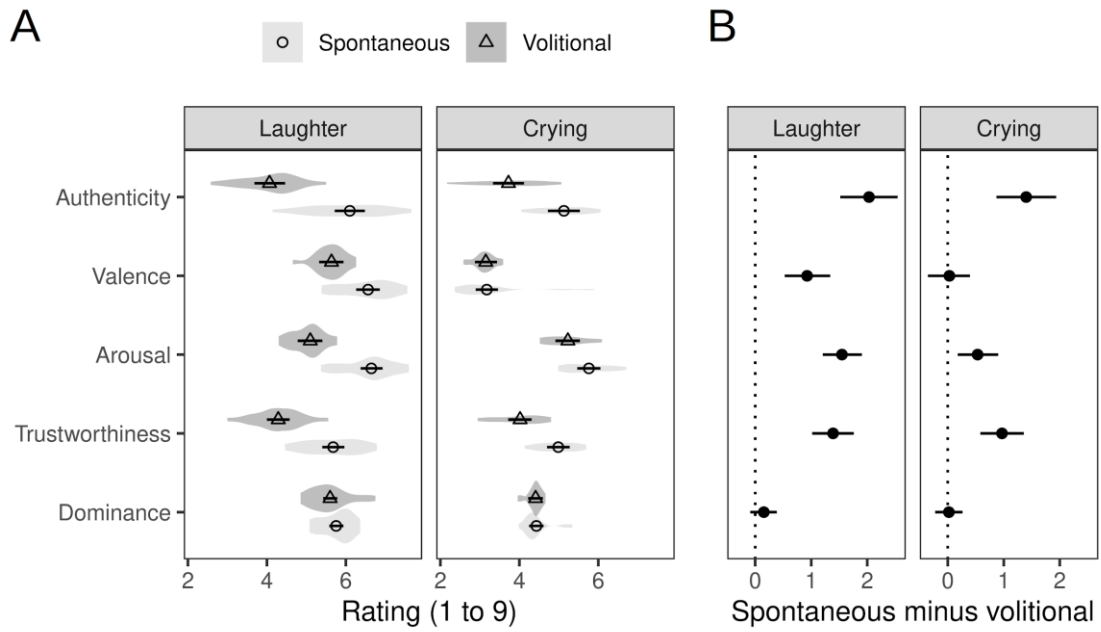


Figure 1. Ratings on each of five scales as a function of the spontaneous or volitional nature of the rated vocalizations (A) and the difference in ratings between spontaneous and volitional vocalizations (B). The points and error bars show fitted values (medians of posterior distribution with 95% CI), and violin plots show the distribution of mean observed values per stimulus.

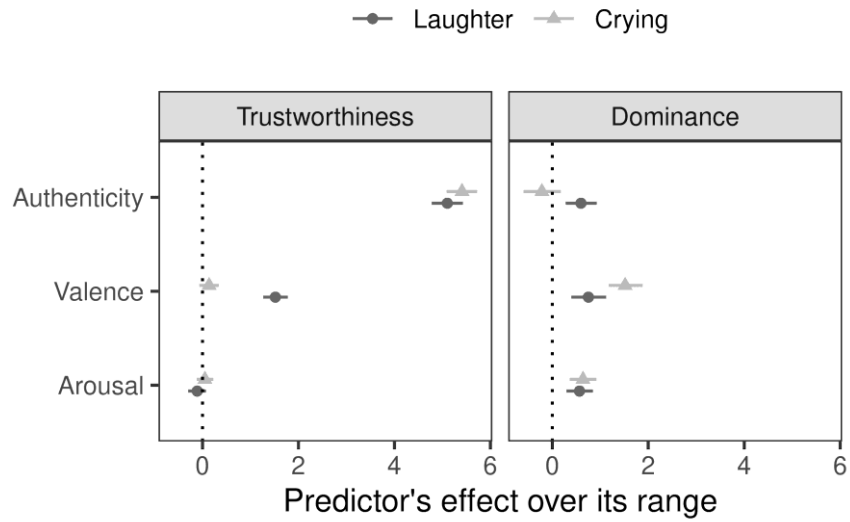
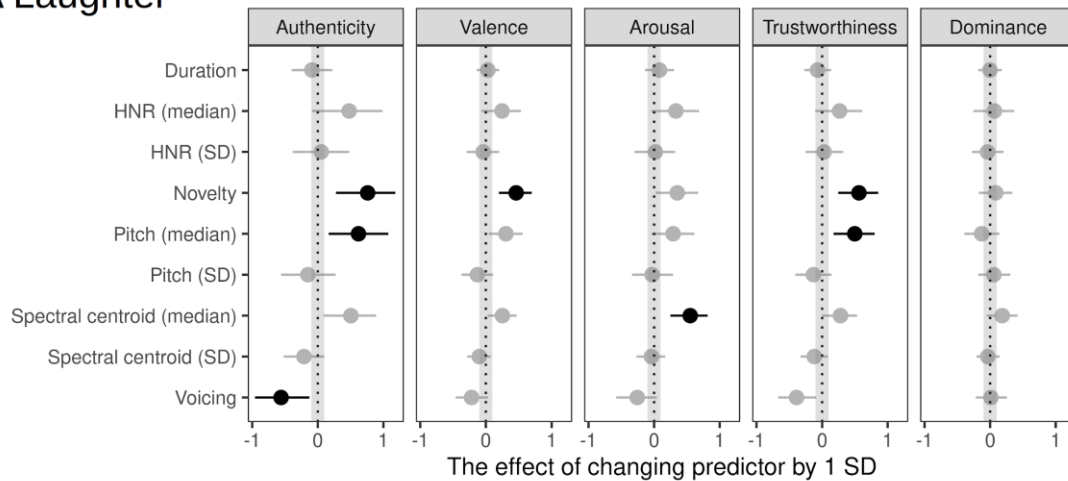


Figure 2. Effects of authenticity, valence, and arousal ratings on the perceived trustworthiness and dominance of the speaker, on a 1-9 scale in multiple regression. Medians of posterior distributions and 95% CIs.

A Laughter



B Crying

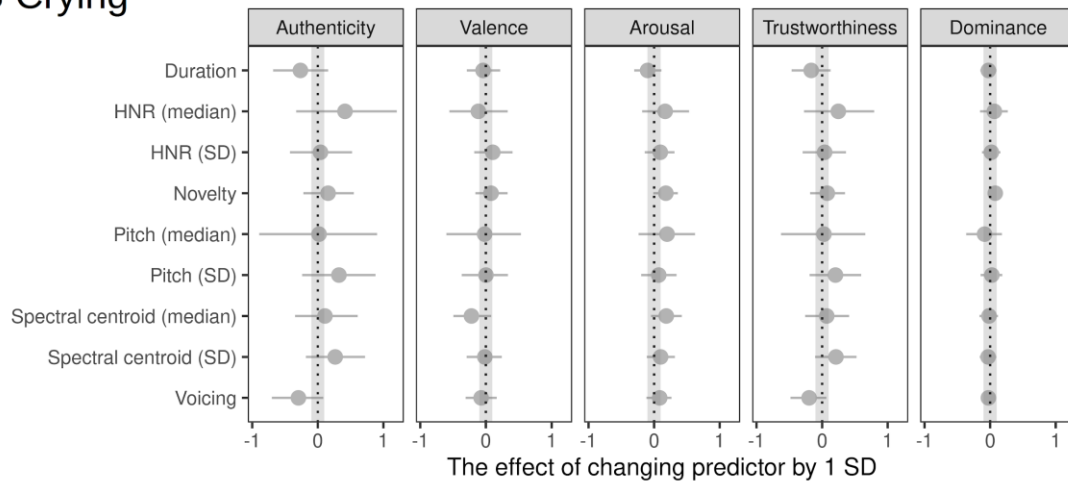


Figure 3. Predicting ratings on the five scales from acoustic characteristics: medians of posterior distribution and 95% CIs from multivariate ordinal regression. The highlighted effects clear the Region of Practical Equivalence (ROPE) of (-0.1, 0.1), corresponding to an effect of at least 0.1 on the rating scale (1 to 9) when changing the predictor by 1 SD. HNR = harmonics-to-noise ratio.