Real-time magnetic resonance imaging reveals distinct vocal tract configurations during spontaneous and volitional laughter

Michel Belyk[1,2] & Carolyn McGettigan[2]

[1] Department of Psychology, Edge Hill University, Ormskirk, United Kingdom

[2] Department of Speech Hearing and Phonetic Sciences, University College London, London, United Kingdom

RUNNING TITLE: Laughter rtMRI

Keywords: Spontaneous laughter, Volitionary laughter, fake laughter, real time magnetic resonance imaging, rtMRI, vocal tract

Correspondence to:
Carolyn McGettigan, Ph.D.
Department of Speech, Hearing and Phonetic Sciences
University College London
London, United Kingdoms
e-mail: c.mcgettigan@ucl.ac.uk

Abstract

A substantial body of acoustic and behavioural evidence points to the existence of two broad categories of laughter in humans: spontaneous laughter that is emotionally genuine and somewhat involuntary, and volitional laughter that is produced on demand. In this study, we tested the hypothesis these are also physiologically distinct vocalisations, by measuring and comparing them using real-time MRI (rtMRI) of the vocal tract. Following Ruch & Ekman (2001), we further predicted that spontaneous laughter should be relatively less speech-like (i.e. less articulate) than volitional laughter. We collected rtMRI data from five adult human participants during spontaneous laughter, volitional laughter, and spoken vowels. We report distinguishable vocal tract shapes during the vocalic portions of these three vocalisation types, where volitional laughs were intermediate between spontaneous laughs and vowels. Inspection of local features within the vocal tract across the different vocalisation types offers some additional support for Ruch and Ekman's (2001) predictions. We discuss our findings in light of a dual-pathway hypothesis for the neural control of human volitional and spontaneous vocal behaviours, identifying tongue shape and velum lowering as potential biomarkers of spontaneous laughter to be investigated in future research.

Introduction

Human laughter offers a unique window into the evolution of vocal behaviour (Pisanski et al., 2016), because it is observed as both a basic emotional vocalisation (spontaneous laughter), and as a highly controlled emotional expression that can be deployed in nuanced ways during social interactions (volitional laughter; Scott et al., 2014). In line with a dual pathway account of the neural control of the human voice (Jürgens, 2002), it is suggested that spontaneous laughs are generated via an evolutionarily conserved neural pathway in the brain's midline, while volitional laughs are controlled by a human-specific neural pathway originating in lateral motor cortex that supports the production of learned vocalisations such as speech and song (Ruch & Ekman, 2001; Wild et al., 2003).

The notion of spontaneous and volitional laughter as distinct vocalisations is supported by a wealth of research on the acoustics and perception of human laughter vocalisations. Although both spontaneous and volitional laughter exhibit a characteristic pattern of repeating "bursts" or "calls" of unvoiced[1] exhalation followed by a vowel-like portion (i.e. the classic "ha ha ha" form), spontaneous laughs have, for example, been reported to be higher in fundamental frequency (f0[2]), be longer in overall duration, and have more (frequent) unvoiced portions (Bryant & Aktipis, 2014; Lavan et al. 2016). Indeed, spontaneous laughs can be confusable with animal vocalisations under certain conditions, supporting the notion that these arise from an older vocal control system shared across apes (Bryant & Aktipis, 2014). Further, these types of laughter communicate perceptually distinguishable social and emotional cues to human listeners: Listeners typically show above-chance accuracy in classifying spontaneous and volitional laughs as, for example, "real" versus "posed" (Bryant & Aktipis, 2014; Bryant et al., 2018; Lavan et al. 2016; Lavan & McGettigan, 2017; McGettigan et al., 2015). Spontaneous and volitional laughs also appear to differentially encode information about talker identity – even when laughs are matched for perceived arousal, listeners' accuracy in voice identity discrimination is lower when listening to laughs that are produced spontaneously (Lavan et al., 2018). These studies all point toward the possibility that spontaneous and volitional laughs may differ in a fundamental sense.

Neurological and neuroscientific investigations have provided additional evidence addressing the hypothesised difference between the neural generators of spontaneous and volitional laughter types. Wild et al. (2003) describe lesion evidence suggesting a double dissociation

---

[1] Voicing describes the articulatory state of the vocal folds in the larynx; *voiced* sounds are made when the vocal folds are held together and are caused to vibrate as air passes through them *en route* from the lungs. In contrast, *unvoiced* sounds are made when the vocal folds are held apart. The difference between a voiced and an unvoiced speech sound can be detected in the difference between the sounds at the start of "bin" and "pin", where the former is voiced and the latter is unvoiced.

[2] The fundamental frequency (F0) is related to the rate of vibration of the vocal folds and is discernible as the perceived pitch of a voiced sound, where higher rates of vibration are related to higher apparent pitch.

between the ability to produce facial expressions (e.g. smiling) volitionally and the spontaneous performance of the same expressions. They implicate subcortical and brainstem structures in the generation of emotional laughter, and lateral motor cortical areas in both the *inhibition* of emotional laughter and in laughing volitionally. More recent studies have elaborated upon this using intracranial electrical stimulation in pre-surgical epilepsy patients. These have implicated the anterior cingulate cortex (ACC) in triggering both affective and motoric aspects of laughter, while the frontal operculum (a lateral motor cortical region) was less reliably associated with mirth. Tractography of human MRI data further suggested differential roles for the frontal operculum and ACC on the basis of their connectivity to other sites implicated in the generation of laughter (Gerbella et al., 2021).

One functional MRI study in healthy participants directly compared task-related neural activation during on-demand volitional laughter production with relatively more involuntary laughter elicited by tickling (Wattendorf et al, 2013). They found that spontaneous laughter was associated with significantly greater activation in the hypothalamus, which Wild et al. (2003) identify as having a key role in laughter generation and affective experience. The ACC was activated during volitional laughter and in the inhibition of ticklish laughter, but not during spontaneous laughter: Although this might appear to contradict findings from Caruana and colleagues (2016), it can be interpreted within the dual pathway model of vocal control proposed by Jürgens (2002). In that account, the ACC is involved in the voluntary initiation (and suppression) of both innate and learned vocalisations, where the former arise via connections to vocal pattern generators in periaqueductal grey to produce innate sounds, and the latter implicate the lateral primary motor cortex in direct connections to brainstem motor nuclei to *shape the content* of learned vocalisations.

Perhaps somewhat surprisingly, in Wattendorf et al.'s (2013) study both spontaneous and volitional laughter as well as laughter inhibition similarly activated a common sensorimotor network including the frontal operculum, the primary motor and somatosensory cortices, and the supplementary motor area (SMA) – thus, the lateral motor control system did not appear to be selectively engaged for voluntary laughter production. However, the experimental context must be taken into account: because excessive head movement leads to artefactual signal in functional MRI data, participants are instructed to minimise movement while being scanned. Thus, in Wattendorf et al's study it becomes difficult to disentangle brain activation due to laughter itself from activation associated with maintaining a steady head position. This conflict, among other experimental constraints, may have masked any true differences in the relative use of the evolutionarily newer lateral motor cortical control pathway for spontaneous and volitional laughter production.

Additional insights on the differences between spontaneous and volitional laughter can be found in the behaviour of the human vocal tract itself during laughter. The human vocal tract – comprising the larynx and the supralaryngeal vocal articulators (e.g. lips, jaw, tongue, velum) – provides the physiological "ground truth" of vocal behaviour, being the physical instrument that gives rise to the sounds of laughter. If spontaneous and volitional laughter are associated with distinct neural systems, it may be possible to see these distinctions within the configurations of the vocal tract during vocalisation (Ruch and Ekman, 2001). Ruch and Ekman (2001) see spontaneous laughter as an involuntary, emotional vocalisation, while volitional ("voluntary") laughter is considered as a controlled behaviour that can be produced independently of a positive emotional experience. They hypothesise that spontaneous laughter's emotional and involuntary nature should manifest in particular effects on both the larynx and the configuration of the articulators within the supralaryngeal vocal tract. Specifically, if spontaneous laughter pre-dates speech, Ruch and Ekman (2001) claim it should be possible to demonstrate that it is an *inarticulate* vocalisation: it should be "generated almost exclusively by laryngeal modulations, modified by some degree by supralaryngeal activity but not by articulation" (p.427). For example, when considering the voiced portions of laughter bursts, a lack of active articulation would predict a relatively central tongue position resemblant of the tongue's resting state during spontaneous laughter, and distinct from the articulated state that gives rise to spoken vowels. However, they note that even an inarticulate tongue may be influenced by the movements of other muscles involuntarily affected by the genuine emotional state in which spontaneous laughter is produced – for example, the opening of the jaw and retraction of the lips in a smile, as well as a widening of the pharynx (the posterior portion of the vocal tract between the velum and the larynx) during positive emotional states.

Ruch and Ekman's (2001) account suggests the vocal tract as a promising locus for the comparison of different laughter types. Magnetic resonance imaging (MRI) offers a way to observe vocal tract behaviour during laughter: Unlike other instrumental methods for the study of speech and articulation, MRI is completely non-invasive and allows the researcher to image the entire vocal tract from the lips to the larynx, at multiple instances per second during vocal behaviour. With its good spatial resolution of anatomical structures, it is possible to obtain global measures of the whole vocal tract in action while maintaining the ability to additionally analyse and interpret local effects (e.g. Belyk et al., 2019; Carey et al., 2017; Carignan et al., 2019, Narayanan et al., 2014; Waters et al., 2021; Wiltshire et al., 2021).

In the current study we therefore used vocal tract MRI (see Figure 1) to empirically compare spontaneous and volitional laughter, and to test Ruch and Ekman's (2001) specific proposals. We used real-time vocal tract MRI to acquire sagittal images of the vocal tract while participants produced spontaneous laughs, volitional laughs, and spoken syllables (e.g. "ha ha ha"). These images were used to trace the outline of the vocal tract during the vocalic portions of individual bursts/syllables – these outlines were then subjected to statistical analysis to describe their multidimensional structure, and to statistically compare this by vocalisation type.

Based on these images, we aimed to empirically test the broad hypothesis that there are physiological differences in the vocal tract the way spontaneous and volitional laughter are produced. We furthermore tested a secondary hypothesis to contextualise the nature of volitional laughter: We reasoned that volitional laughter is generated by the same neural system that produces speech, in order to volitionally simulate laughter in lieu of neural pathways that would generate it spontaneously. We therefore predicted that 1) spontaneous and volitional laughter should be distinguishable in the vocal tract and 2) there should be greater similarity between volitional laughter and speech in the vocal tract than between spontaneous laughter and speech.

Methods

Participants

A total of five adults (4 female, 1 male), completed the study. Participants were recruited from the staff and PhD student population of the Department of Psychology at Royal Holloway, University of London, who were familiar with the research team and environment. This sampling strategy was used to maximise the chance of obtaining samples of spontaneous laughter, as unfamiliar participants may feel more inhibited by the unusual environment of the MRI. For inclusion, participants were required to be aged between 18 and 40, with healthy hearing (self-reported) and no neurological illness (self-reported). The data from a sixth participant was discarded due to technical issue during scanning. This study was approved by the Department of Psychology Ethics Committee at Royal Holloway, University of London and participants provided written informed consent.

Procedure

Participants underwent 4 runs of rtMRI each in which they laughed spontaneously at self-selected humorous videos, laughed volitionally (on demand) while watching non-humorous videos, or spoke canonical vowels in Standard Southern British English. One participant completed 3 runs of spontaneous laughter due to technical difficulties during one run in the presentation of their self-selected videos. Two participants each completed one additional run of spontaneous laughter. Conditions were always completed in the order of vowels, voluntary laughter, then spontaneous laughter, in order to prevent the contamination of the former conditions by spontaneous laughter.

In rtMRI runs of spontaneous laughter, participants were presented with audiovisual clips that they had previously selected as likely to induce audible laughter. Examples of clips included scenes from popular television shows (e.g. *Friends*), feature films, amusing videos of animals, and material related to the participants' individual interests (e.g. the Eurovision Song Contest). Participants produced laughter spontaneously when they found the clips amusing. In runs of volitional laughter production, participants viewed a control clip of a narrated demonstration in the statistical software SPSS (IBM, Armonk, NY), which was selected as an example of non-humorous material ("SPSS for Beginners 1: Introduction" https://www.youtube.com/watch?v=ADDR3_Ng5CA). Participants were instructed to watch the video and produce laughter "on demand" regularly throughout the scan. In vowel runs, participants were provided with an onscreen cue instructing them to repeat one of the syllables "hee", "her", "hoo", "hah", or "har" (/hiː/, /hɜː/, /huː/, /hæ/, /hɑː/). The vowels were selected to provide approximate coverage of the four corners and centre of the English vowel quadrilateral. Each vowel was repeated slowly in blocks of 5 vocalisations, at a rate of approximately 0.5 Hz. This slow rate of articulation ensured that a larger proportion of rtMRI frames would occur on the steady state of the vowel. All stimuli were presented via the Psychophysics toolbox running in Matlab (The Mathworks, Natick, MA). Audio stimulation was delivered through MR-compatible earbuds (Sensimetrics Model S14, Sensimetrics Corporation, Gloucester, MA). Visual stimuli were delivered via back projection of visual stimuli onto an in-bore screen, and viewed via a mirror mounted on the headcoil. Audio vocalisatoin data were recorded inside the scanner using a fibre-optic microphone (FOMRI-III; OptoAcoustics Ltd, Or Yehuda, Israel).

Real-time magnetic resonance imaging

Real-time MRI (rtMRI) data were fast gradient echo images collected on a Siemens 3T TIM Trio scanner; flip angle: 5°; TE/TR: 1.25/3.2 ms; GRAPPA factor 2; partial-Fourier: 75%; FOV 220 × 274 mm^2; 2.5 × 2.5 × 10.0 mm^3 spatial and 125 ms temporal resolution (8 frames per second [f.p.s.]). Although the frame rate is relatively slow compared with those reported in other vocal tract MRI studies (Carignan et al., 2019, Narayanan et al., 2014; Wiltshire et al., 2021), it was sufficient to capture the vocal portions of the behaviours measured in the current experiment. Each rtMRI run spanned 500 frames, to a total of 1500-2500 frames per participant per condition.

Analysis

Identifying vocalisations from in-scanner recordings

Audio recordings were aligned to the onset of rtMRI runs and denoised using Audacity (Audacity Team, 2018) - the spectrum of the MRI scanner noise was estimated from a period during which the participants was silent, then removed by subtraction from the whole audio recording. The onsets and offsets of bouts of vocalisation were semi-automatically identified using an in-house Praat script (Boersma & Weenink, 2019) which identifies silent versus sounding portions of the audio recordings and identifies rtMRI frames within each run that occurred when participants were vocalising. The outcomes of this automatic detection were manually checked and hand corrected by author MB. Speakers produced 449-1309 frames of spontaneous laughter, 339-1077 frames of volitional laughter, and 461-890 frames of vowels (see Table 1).

Vocal tract tracing

We used a custom-built toolbox (Belyk, Carignan, McGettigan, pre-print), which semi-automatically extracts the shape of the vocal tract from rtMRI data using spatially constrained tissue classification. Each rtMRI frame was registered to a common reference image using rigid body transformation. The approximate location of the vocal tract within the rtMRI series was estimated by identifying high variance pixels, since alternation between high intensity (soft tissue) and low intensity (air) is a characteristic of vocal tract pixels. An informed analyst (MB) then manually adjusted this estimate to create a mask that identified pixels that may sometimes contain vocal tract. These pixels were then subject to simple tissue classification based on the high degree of contrast between air and soft tissue in T1-weighted images. The resulting tissue masks were then converted to outlines, manually inspected, and corrected for tissue classification errors where necessary.

Functional principal components analysis

Vocal tract traces were analysed using functional principal components analysis (fPCA) (Ramsay et al., 2009; Ramsay & Silverman, 2005) in R (R Core Team, 2019; Ramsay et al., 2017) following a method we have previously demonstrated on outlines of the tongue during whistling (Belyk et al., 2019). Functional PCA explores patterns of variation in the shapes of functions around a mean shape. Much like discrete PCA, fPCA seeks principal components that maximize variation between observations (Levitin et al., 2007; Locantore et al., 1999). The principal components of discrete PCA are eigenvectors that map each component back onto a set of discrete variables. Similarly, the principal components of functional PCA are eigenfunctions that map each component back onto variations in shape. Applied to the two-dimensional coordinates of the outline of the vocal tract, this approach provides an empirical means of studying changes in vocal tract shape.

Selection of vocal tract shapes for analysis: Identifying steady-state portions of vocalic behaviours

An initial fPCA identified frames associated with steady-state vocalic portions of the utterances. This initial analysis revealed that vocal tract shapes fell into two discrete clusters, based primarily on fPC1. These two clusters were further isolated using K-means clustering based on the first four fPCs. Cluster 1 was the smaller of these clusters and consisted of 2786 rtMRI frames that overwhelmingly occurred at the onset or the offset of bouts of vocalisation, with few exceptions. Cluster 2 was the larger of these clusters and consisted of 8568 rtMRI frames that were associated with the central portion of the vocalisation during which the vocal tract is expected to reach a steady state. Consistent with this interpretation, positive scores on fPC1 (associated with Cluster 1) indicated consonant-like constriction of the vocal tract at the velum or the palate, while negative scores on fPC1 (associated with Cluster 2) indicated a vowel-like configuration of the vocal tract which remains unconstricted throughout. A subsequent fPCA was therefore conducted using only the steady state frames identified by Cluster 2 membership, and further analyses are restricted to these data (see Supplementary Materials 1). The final analysis included 253-855 frames of spontaneous laughter, 425-1029 frames of volitional laughter, and 173-759 frames of vowels per participant (see Table 2).

Results

Qualitative description

A qualitative view of the mean vocal tract shape during spontaneous laughter, volitional laughter, and vowels (see Figure 2A; note that for illustration vocal tracts are shown with the origin centred at the aperture of the lips, as this point could be reliably identified by automatic processes) suggests that volitional laughter was produced with a vocal tract shape that was intermediate to spontaneous laughter and vowels. Spontaneous laughter was associated with a longer overall vocal tract outline suggestive of a lowered larynx, an overall flatter and less bunched tongue position, and greater constriction of the vocal tract around the velum suggestive of velum lowering. This overall pattern was relatively consistent across participants (Figure 2B).

Functional principal components analysis

A qualitative accounting of vocal tract shape alone does not account for the potentially large degree in variation within vocalisation types. The techniques of functional data analysis Ramsay et al., 2009; Ramsay & Silverman, 2005) provide a robust framework with which to quantify variation in vocal tract shape.

Functional principal components analysis identified a small number of components which described the principal modes of variation in the shapes of the vocal tract. An examination of the scree plot for this analysis (See Supplementary Materials 1) revealed that the first four functional principal components (fPCs) explained greater than 80% of vocal tract shape variation, and that the explanatory value of examining further components diminished rapidly.

Each functional principal component reflects complex variation, affecting several aspects of the vocal tract outline. While we provide subjective descriptions of each component, we caution that fPCA is data driven and not biologically constrained. Furthermore, vocal tract visualisations are shown with the origin centred at the aperture of the lips to provide a common space for comparison, which may induce small variations in the position of the image origin both within and between participants. Therefore, vocal tract outlines will be affected not only by the behaviour, but also by between-person variation in anatomy. Hence, descriptive accounts of individual fPCs must be treated with caution.

- **The first component (fPC1):** *Tongue bunching*: This component describes variation from a bunched and anterior tongue configuration for negative scores, to a slightly backed and flatter configuration for positive scores. The vertical position of the larynx is higher than the mean at low fPC scores and lower than the mean at higher fPC scores. Inspection of the fPC values by vocalisation type suggests scores around zero for the majority of spontaneous laughter frames, with vocal tract configurations similar to the overall mean, while volitional laughter and vowels have negative scores (see Figure 3A).
- **The second component (fPC2):** *Tongue backing and tract curvature*: This component ranges from a slightly fronted tongue for negative scores to a more backed tongue for positive scores. However, there is also variation in overall vocal tract shape:

negative scores reflect a lower larynx and greater tract curvature posterior to the velum. Spontaneous laughs load more negatively on this component than both volitional laughs and vowels, where vowels have the most positive weightings (see Figure 3B).

- **The third component (fPC3):** *Velum raising and lowering*: This component ranges from a narrowed/constricted vocal tract at the velum for negative scores, to a wider velar aperture for positive scores. Spontaneous laughs load more negatively on this component than volitional laughs, which in turn are weighted less positively than vowels (see Figure 4A).
- **The fourth fPC (fPC4):** *Tongue shape and height:* This component ranges from low and flat tongue shape with pharyngeal constriction, to high and bunched tongue shape with slight pharyngeal widening. Vowels tend to show more positive scores than laughter, where spontaneous and volitional laughter show similar overall scores. However, laughter is only associated with negative scores in some of the participants (see Figure 4B).

As for the average vocal tract outlines described above, plots of the fPC values for individual analysis frames show a relatively consistent pattern across participants, where volitional laughs lie intermediate between spontaneous laughs and vowels (see Figures 3C and 4C). An RShiny companion app to this article provides interactive visualisation and data exploration from these functional principal components individually or in combination (see Figure 5)

Euclidean distances between vocalisation types in fPC space

This analysis aimed to establish whether clusters of spontaneous laughs, volitional laughs, and vowels were distinct within the multidimensional fPC space. Information was combined across fPCs by computing the Euclidean distance from each rtMRI frame to each of the run-type centroids (i.e., the distance to the centroid of each of spontaneous laughter, volitional laughter, and vowels), for each speaker. Euclidean distances were modelled using a linear mixed model (see Supplementary Materials 3 for model structure and diagnostics), from which were derived estimates and confidence intervals of the Euclidean distance of each vocalisation type to its own category centroid and to the centroids of each other category of vocalisation (see Figure 6).

Each category of vocalisation had smaller distances to its own centroid than to the other group centroids, indicating that spontaneous laughter, volitional laughter, and vowels were distinguishable as vocalisation categories based solely on the shape of the vocal tract. Moreover, the Euclidean distance between volitional laughter and the other two categories was smaller than the distance between spontaneous laughter and vowels. The vocal tract shape of volitional laughter was therefore intermediate between spontaneous laughter and speaking isolated vowels in the multidimensional space defined by the 4 fPCs.

Univariate analyses of individual fPC scores

Scatterplots of fPC scores (see Figure 3C and Figure 4C) demonstrate that the vocal tract shapes of spontaneous laughter, volitional laughter, and vowels are distinguished multivariately by combinations of fPCs more than by any one component in isolation. Regardless, it can be informative to try to understand the contribution of each component to distinguishing between each category of vocalisation. Linear mixed models were computed separately predicting each of fPCs 1-4 from a fixed effect of run type and random slope of run type within speaker (see Supplementary Materials 3). The interpretation of these analyses should be tempered by the relatively small number of speakers contributing to each model.

In the results that follow, spontaneous laughter was modelled as the reference category and contrast estimates are provided against volitional laughter and vowels. The first fPC did not significantly distinguish between vocalisation categories (F(2, 4) = 4.38, p = 0.098), although there were marginal differences from spontaneous laughter (Volitional: estimate = -7.3, t(4) = -2.51, p = 0.066; Vowels: estimate = -7.2, t(4) = -2.51, p = 0.066). The second fPC significantly distinguished between vocalisation categories (F(2, 4) = 29.9, p = 0.0038) and this was primarily driven by differences between spontaneous laughter and vowels (Volitional: estimate= 2.2, t(4) = 1.1, p = 0.32: Vowels: estimate = 10.2, t(4) = 4.2, p = 0.014). The third fPC also distinguished between vocalisation categories (F(2, 4) = 20.1, p = 0.0081), where spontaneous laughter was significantly different from both volitional laughter and vowels (Volitional: estimate = 3.2, t(4) = 5.3, p = 0.006; Vowel: estimate = 5.8, t(4) = 3.3, p = 0.03). The fourth fPC displayed little to no explanatory value (F(2, 4) = 0.006, p = 0.99; Volitional: estimate = -0.16, t(4) = -0.10, p = 0.93; Vowel: estimate = -0.063, t(4) = -0.03, p = 0.98). Together these findings suggest that spontaneous laughter is distinct from vowels in larynx height and tongue backness (fPC2), while also showing greater velar lowering than both volitional laughter and vowels (fPC3).

Discussion

This study tested the hypothesis that spontaneous and volitional laughter are two distinct vocalisation types, which may be controlled by two different neural pathways in the human brain. We compared the vocal tract shapes of five human participants while they produced spontaneous and volitional laughs, and spoken vowels. Our specific predictions were that vocal tract configurations during spontaneous and volitional laughter should be distinct, and that volitional laughs should have greater similarity to vowels.

We found supportive evidence for our hypotheses across qualitative and quantitative examinations of vocal tract shapes: the properties of the vocal tract during volitional laughter were intermediate between those of spontaneous laughter and vowels, and the distances between vocalisation types showed greater similarity between volitional laughter and vowels than between spontaneous laughter and vowels. This relationship between vocalisation types – seen at the level of individual participants as well as the group – is compatible with an interpretation of volitional laughter as being relatively more similar to speech compared to spontaneous laughter. When humans laugh volitionally, we suggest that they are using the neural pathway associated with speech motor control to mimic the sounds of laughter in its spontaneous forms. This ability to simulate a spontaneous vocalisation, albeit imperfectly, may be adaptive – signalling positive emotion even in the absence of genuine emotional experience may facilitate the formation of interpersonal social bonds and advance the laugher's admission to social groups (Bryant & Aktipis, 2014; Curran et al., 2015).

Notably, all exemplar frames were treated equally in the fPC analysis, and yielded clearly distinct clusters associated with each vocalisation type. Our analysis of the Euclidean distance between individual vocalisation frames and their category centroids further supports the validity of spontaneous and volitional laughter as distinct types of vocalisation: laughs generated spontaneously during genuine amusement in our study are more similar to other laughs generated in this same state than to volitional laughs generated "on demand" (and vice versa). It is important to note that laughter frames were not chosen for analyses based on any prior perceptual validation in terms of their discriminability or perceived authenticity – all frames that were viable for analysis were included and labelled only according to the context in which they were produced, not on the basis of whether they sounded sufficiently "real" or "posed". Thus we interpret the findings on the basis that laughs produced spontaneously are different from those that are produced volitionally. This echoes previous findings in perception studies – Lavan and colleagues (2018) found that listeners were less accurate at discriminating voice identity from spontaneous laughter than volitional laughter, suggesting that spontaneous laughter is a distinct type of vocal act in which indexical person characteristics are more poorly encoded.

Spontaneous laughs showed a flatter and lower tongue configuration, a longer overall vocal tract outline consistent with larynx lowering, and relatively greater constriction around the velum suggestive of velum lowering. Ruch and Ekman (2001) proposed that spontaneous laughter should resemble an “inarticulate” vocalisation. With the caveat that the effects of gravity due to the supine position of our participants will affect tongue shape overall due to the effects of gravity pulling the tongue toward the back of the throat, the average outlines of the vocal tract in Figure 2 indicate a relatively flatter and less bunched tongue configuration in spontaneous laughter, relative to volitional laughter and vowels. Within the fPCA, variation in tongue shape and position is seen most clearly along fPC1, fPC2, and fPC4. In both fPC1 and fPC2, it is striking that the weightings for spontaneous laughter tend to implicate a tongue position and shape that overlaps with the grand mean vocal tract outline, which may suggest a somewhat inarticulate tongue as suggested by Ruch and Ekman. However, of the tongue-related components the only statistical difference between spontaneous laughs and vowels was found on fPC2, which additionally implicated vocal tract lengthening (i.e., larynx lowering) in spontaneous laughs and shortening in vowels. We note that fPC2 also carries some variation suggestive of overall changes in vocal tract curvature, which implies the contribution of between-subject variations in vocal tract anatomy.

Ruch and Ekman (2001) also consider the role of the velum (or soft palate) in their discussion of the supralaryngeal articulators in laughter. It is not clear whether the neutral state of the velum during vocalisation is to be closed, thus diverting all respiratory airflow through the oral cavity, or open (partially or fully) and diverting air through the nasal cavity. In speech, the presence of *nasality* is associated with an increase in low-frequency acoustic energy, and our own previous work on the perception of laughter found an increased perception of nasality and reduced perception of mouth-opening with low-authenticity volitional laughs (Lavan et al., 2016). However, if we consider the state of the velum during rest, it is necessarily lowered to allow aerobic respiration to continue when the mouth is shut. In the current study it is spontaneous laughter that appears to exhibit a more lowered velum, indicated by greater constriction in this portion of the vocal tract outline. This is also shown in fPC3 of the functional PCA, where we also found statistically significant differences in the component weightings between spontaneous laughter and both volitional laughter and vowels. In line with Ruch and Ekman’s proposal that the supralaryngeal articulators should be in their resting position during spontaneous laughter, a lowered velum would indeed be the inarticulate state of this structure, outside of speech.

Another proposed substrate for differences between spontaneous and volitional laughter was in the width of the pharyngeal portion of the vocal tract, between the velum and the larynx. Although there was some apparent variation in pharyngeal width across the fPCs, the interpretability of these effects was limited by several factors. For example, some fPCs showed variation indicative of between-talker differences in vocal tract shape (e.g. overall vocal tract curvature) as well as possible within-talker variation that could be attributed to behaviour. Furthermore, where variation in pharynx width was apparent (e.g. in fPC4), there were no statistical differences in the weighting of the three vocalisation types on this component.

There are several limitations of the study that should be noted. First, the overall participant sample was small, with usable data from only 5 participants. The process of tracking and manually correcting thousands of rtMRI frames is labour-intensive, despite the level of built-in automaticity to our analysis pipeline. On the one hand, the plots of individual participant data show relatively consistent evidence for within-subject separation of the three vocalisation types along the fPCs. But there is also evidence for considerable between-subject variability in the nature of vocal tract configurations by vocalisation type, which may suggest subtle individual differences in the underlying behaviours. A second limitation is that it is difficult to confirm the ground truth of the emotional state of the participants. Genuine emotional experiences cannot be guaranteed, and there may have been variation in the degree to which participants experienced amusement during the spontaneous laughter runs that might introduce heterogeneity in the vocal tract samples. Obtaining perceptual ratings of audio laughter samples could help to determine variation in perceived emotional arousal and authenticity, but this is indirect and furthermore agnostic to the true emotional state of the vocaliser. Future work could therefore seek to obtain self-report measures of emotional state during laughter runs in order to identify changes in the vocal tract that are dependent on the intensity of affective experience. Finally, we must acknowledge that collecting vocal tract MRI data from supine participants limits the generalisability of the precise vocal tract properties we observed, as in everyday life most vocal behaviour is performed when the body is upright. Thus we again note caution in the interpretation of our data with regard to claims about the “inarticulate” states of different articulators.

Conclusions

We have used vocal tract MRI during laughter and spoken vowel vocalisations to examine how spontaneous and volitional laughter manifest in the shape of the human vocal tract, and how these in turn relate to speech. In line with the existing acoustic, perceptual, and neurological evidence, we see consistent evidence for their physiological separability, both across and within participants. Volitional laughs were produced with vocal tract shapes that are intermediate between spontaneous laughter and vowels. This, coupled with indications of reduced articulatory activity in spontaneous laughs (i.e., resting tongue configuration and lowered velum) may support to the hypothesis that spontaneous and volitional laughs are controlled by distinct neural pathways in the human brain – one that is seen in other primates and generates innate emotional vocalisations, and another that is seen most prominently in humans and is associated with learned vocalisations. Without accompanying neural activation data it is impossible to draw conclusions about the relationship between these vocal tract configurations and the brain systems generating them. However, we have presented a starting point for more comprehensive modelling of laughter that incorporates the physiological effects of emotion on the vocal anatomy. Immediate next steps should attempt to replicate our findings in a larger sample and to take steps to ensure, or at least monitor, the level of authentic emotional experience during the production of spontaneous and volitional laughs. Beyond this, it will be important to link vocal tract configurations to patterns of underlying neural activation during laughter. Complementary work could probe the vocal tract shape in individuals who are trained to produce on-demand laughter that connotes greater authenticity – for example, actors and voice artists. Data from such vocal experts could be used to test whether vocal tract shapes during emotionally convincing volitional laughter overlap more closely with those seen during spontaneous laughter, and thus provide evidence on whether authentic emotional experience can indeed be “faked” (Bryant & Aktipis, 2014; McKeown, Sneddon, Curran, 2015). Finally, this study took a broad approach in categorising laughs as broadly spontaneous or volitional, though we appreciate that human laughter is more nuanced and context-dependent than the two versions presented here (Curran et al., 2018) – future investigations should interrogate the vocal tract during laughter that is more reflective of the varied naturalistic social settings in which it is typically observed.

Acknowledgements

This research was supported a Research Leadership Award from The Leverhulme Trust (RL-2016-013; awarded to CM).

Tables

Speaker	Spontaneous Laughter	Volitional Laughter	Vowel
P1	1130	660	890
P2	449	490	461
P3	1309	339	738
P4	1013	651	474
P5	1149	1077	524

Table 1: Summary of the number of frames recorded from each participant and each condition.

Speaker	**Spontaneous Laughter**	**Volitional Laughter**	**Vowel**
P1	387	931	759
P2	253	425	416
P3	278	1029	546
P4	526	742	173
P5	855	746	502

Table 2: Summary of the number of frames from each participant and each condition included in the final analysis.

Figures

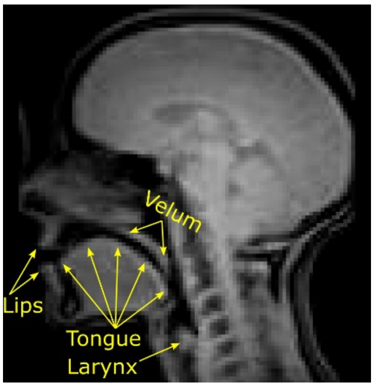


Figure 1: Representative midsagittal image of the vocal tract. T1-weighted images provides contrast between soft-tissue (light) relative to bone and air (dark). The labile structures that shape the vocal tract are labelled (yellow) and the vocal tract itself is composed of the negative space between them.

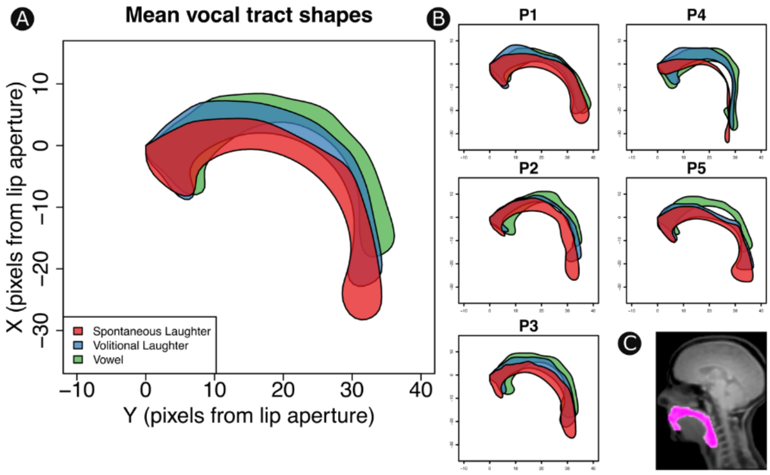


Figure 2: A) Mean vocal tract shapes for spontaneous laughter (red), volitional laughter (blue), and isolated vowels (green). Vocal tracts are shown with the origin centred at the aperture of the lips as this point could be reliably identified by automatic processes. B) Mean vocal tract shapes for each individual speaker. In all cases the vocal tract shape of volitional laughter is intermediate between spontaneous laughter and vowels. C) A representative vocal tract (pink) overlayed with a midsagittal MRI frame for anatomical context.

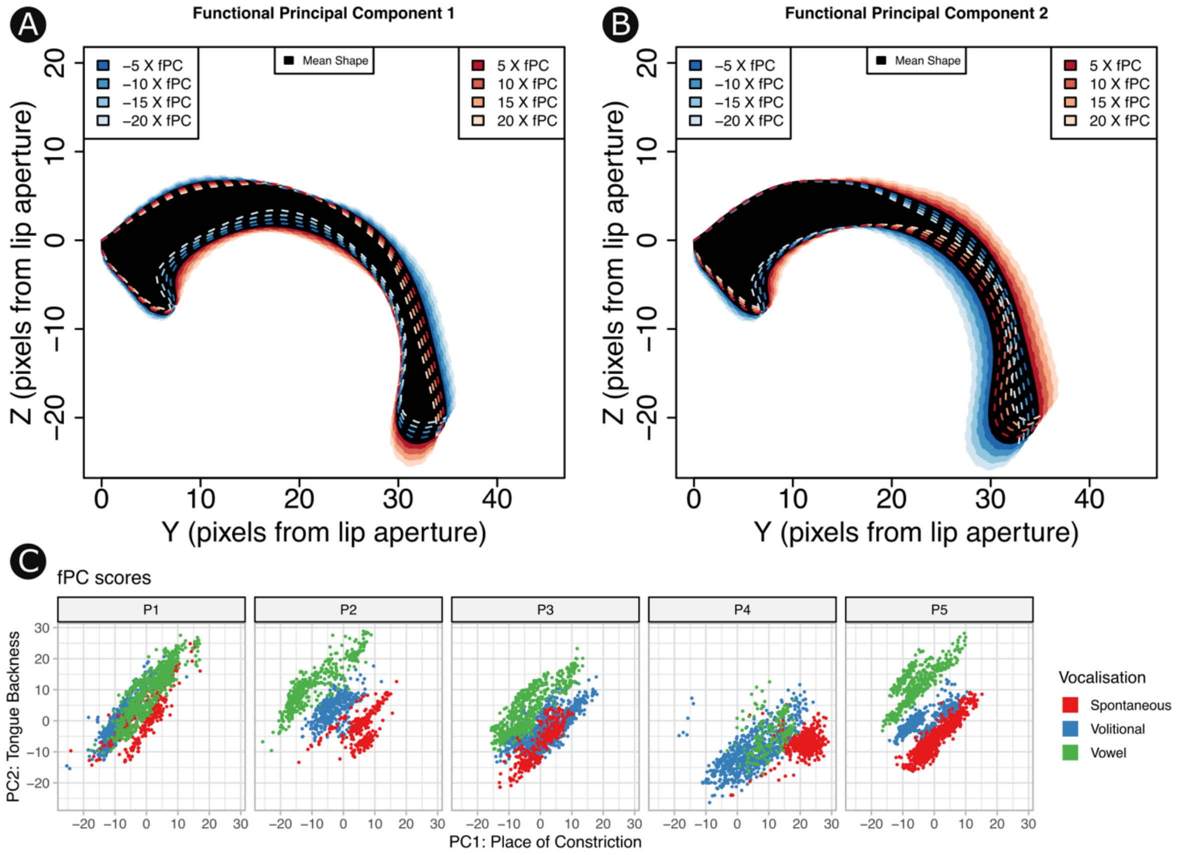


Figure 3: Summary of first and second functional principal components (fPCs) 1 and 2. A) Visualisation of fPC1 accounting for 34.4% of variation of vocal tract shape. The black area depicts the mean shape of the vocal tract; red shading and lines indicate the vocal tract shapes that correspond to increasing fPC1 scores, while blue shading and lines indicate the vocal tract shapes that correspond to decreasing fPC1 scores. B) Visualisation of fPC2 accounting for 30.2% of variation on vocal tract shape. C) Scatterplots of fPC1 and fPC2 scores for each speaker (panel) and each vocalisation category (colour). Each point represents a single imaging frame. An RShiny companion app provides interactive visualisation and data exploration from these functional principal components individually or in combination.

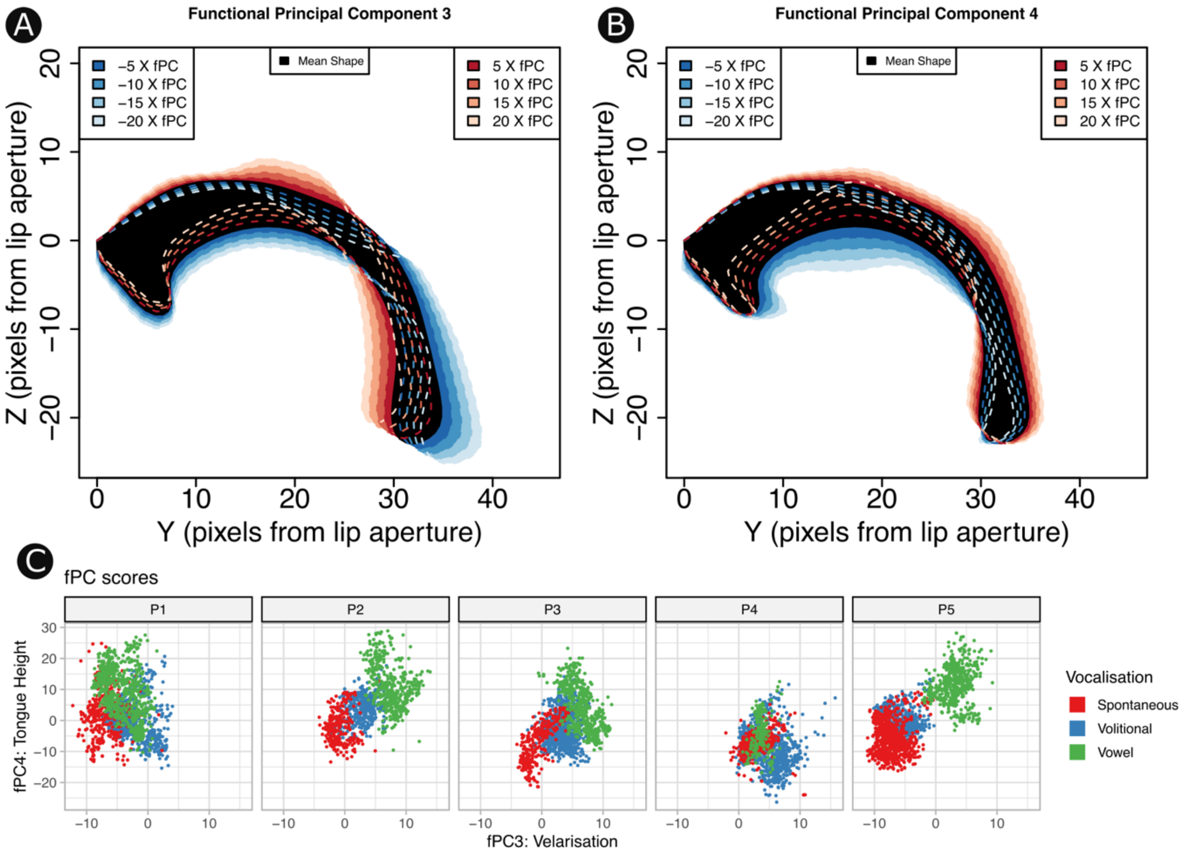
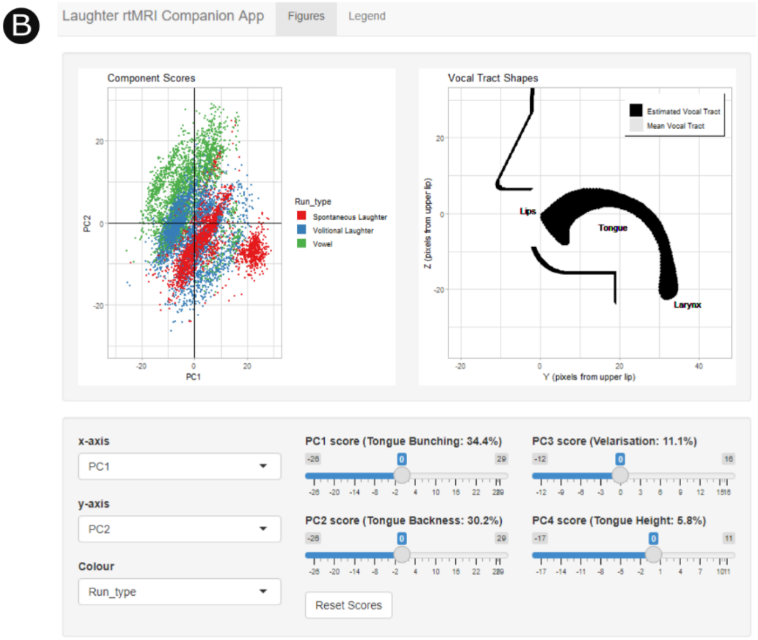


Figure 4: Summary of first and second functional principal components (fPCs) 3 and 4. A) Visualisation of fPC3, which accounted for 11.1% of variation in vocal tract shape. The black area depicts the mean shape of the vocal tract; red shading and lines indicate the vocal tract shapes that correspond to increasing fPC3 scores, while blue shading and lines indicate the vocal tract shapes that correspond to decreasing fPC3 scores. B) Visualisation of fPC4, which accounted for 5.8% of variation in vocal tract shape. C) Scatterplots of fPC3 and fPC4 scores for each speaker (panel) and each vocalisation category (colour). Each point represents a single imaging frame.



x-axis
PC1

y-axis
PC2

Colour
Run_type

PC1 score (Tongue Bunching: 34.4%)
0

PC2 score (Tongue Backness: 30.2%)
0

PC3 score (Velarisation: 11.1%)
0

PC4 score (Tongue Height: 5.8%)
0

Figure 5: This article is accompanied by an interactive data visualisation app with which the reader can explore the first four functional principal components of vocal tract shape during spontaneous laughter, volitional laughter, and vowels. A) The app can be accessed via QR code, url (https://michelbelyk.shinyapps.io/rtMRI_Laughter/), or by downloading the source code and data provided in Supplementary Materials 2. B) Still capture from the app. The scatterplot (top left) shows scores for each vocal tract image and a crosshair to highlight the currently selected combination of principal component scores. The shape plot (top right) shows the corresponding vocal tract shape as well the mean vocal tract shape for comparison. Sliders spanning the range of observed scores in the data are used to dynamically explore changes in the shape of the vocal tract. In the still capture, all components are set to zero which models the mean shape of the vocal tract.

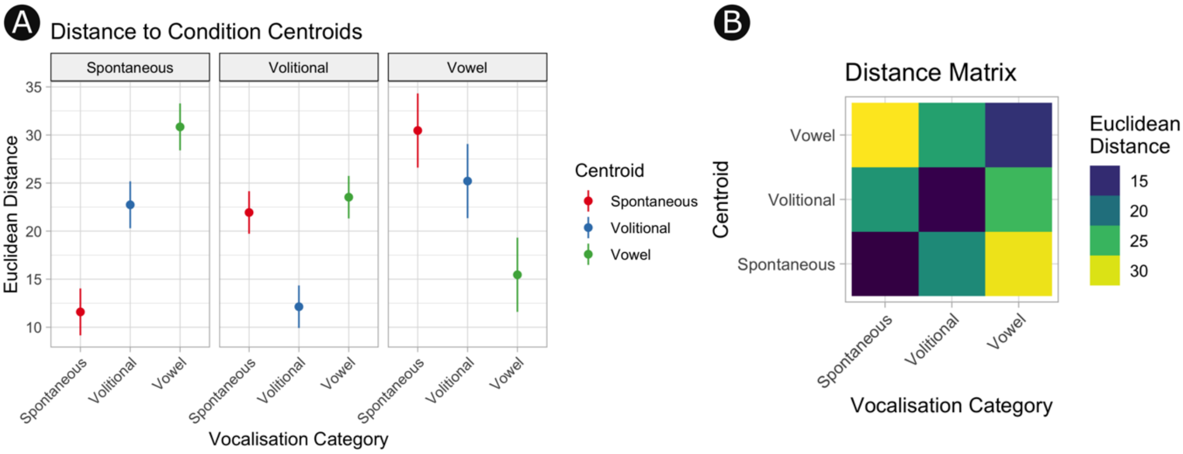


Figure 6: Dissimilarity between vocalisation categories. A) Each panel summarises Euclidean distances from frames of one category of vocalisation (panel title) to vocalisation category centroids (colour). Vocalisations had the least distance to their own category centroid relative to out of category centroids. Volitional laughter was intermediate between spontaneous laughter and vowels. B) Euclidean distances presented in distance matrix form. Each cell depicts the estimated Euclidean distance from one category of vocalisation (x-axis) to one vocalisation category centroid (y-axis). The diagonal reflects distances to within-category centroids. The larger internal distances within the vowel category (top right) reflects the use of a diverse range of vowels in these vocalisations. Off-diagonal cells reflect distances to out-of-category centroids, the greatest of which is between spontaneous laughter and vowels.

References

Audacity Team. (2018). *Audacity(R): Free audio editor and recorder* (2.1.3).

Belyk, M., Carignan, C., & McGettigan, C. (under review). Semi-automatic method for analysing vocal tract morphology from real-time magnetic resonance images. *PsyArXiv* doi: 10.31234/osf.io/y6xde

Belyk, M., Schultz, B. G., Correia, J., Beal, D. S., & Kotz, S. A. (2019). Whistling shares a common tongue with speech: bioacoustics from real-time MRI of the human vocal tract. Proceedings of the Royal Society B, 286(1911), 20191116.

Belyk, M., Schultz, B. G., Correia, J., Beal, D. S., & Kotz, S. A. (2019). Whistling shares a common tongue with speech : Bioacoustics from real-time MRI of the human vocal tract. *Proceedings of the Royal Society B*, *286*, 20191116. https://doi.org/10.1098/rspb.2019.1116

Boersma, P., & Weenink, D. (2019). *Praat: Doing phonetics by computer* (6.1.06). http://www.praat.org/

Bryant, G. A., & Aktipis, C. A. (2014). The animal nature of spontaneous human laughter. Evolution and Human Behavior, 35(4), 327-335.

Bryant, G. A., Fessler, D. M., Fusaroli, R., Clint, E., Amir, D., Chávez, B., ... & Zhou, Y. (2018). The perception of spontaneous and volitional laughter across 21 societies. Psychological science, 29(9), 1515-1525.

Carey, D., Miquel, M. E., Evans, B. G., Adank, P., & McGettigan, C. (2017). Vocal tract images reveal neural representations of sensorimotor transformation during speech imitation. Cerebral Cortex, 27(5), 3064-3079.

Carignan, C., Hoole, P., Kunay, E., Pouplier, M., Joseph, A., Voit, D., ... & Harrington, J. (2020). Analyzing speech in both time and space: Generalized additive mixed models can uncover systematic patterns of variation in vocal tract shape in real-time MRI. Laboratory Phonology: Journal of the Association for Laboratory Phonology, 11(1).

Caruana, F., Gozzo, F., Pelliccia, V., Cossu, M., & Avanzini, P. (2016). Smile and laughter elicited by electrical stimulation of the frontal operculum. *Neuropsychologia*, *89*, 364-370.

Caruana, F., Avanzini, P., Gozzo, F., Francione, S., Cardinale, F., & Rizzolatti, G. (2015). Mirth and laughter elicited by electrical stimulation of the human anterior cingulate cortex. *Cortex*, *71*, 323-331.

Curran, W., McKeown, G. J., Rychlowska, M., André, E., Wagner, J., & Lingenfelser, F. (2018). Social context disambiguates the interpretation of laughter. Frontiers in Psychology, 8, 2342.

Gerbella, M., Pinardi, C., Di Cesare, G., Rizzolatti, G., & Caruana, F. (2021). Two neural networks for laughter: A tractography study. *Cerebral Cortex*, *31*(2), 899-916.

Lavan, N., Short, B., Wilding, A., & McGettigan, C. (2018). Impoverished encoding of speaker identity in spontaneous laughter. *Evolution and Human Behavior*, *39*(1), 139-145.

Lavan, N., & McGettigan, C. (2017). Increased discriminability of authenticity from multimodal laughter is driven by auditory information. Quarterly Journal of Experimental Psychology, 70(10), 2159-2168.

Levitin, D. J., Nuzzo, R. L., Vines, B. W., & Ramsay, J. O. (2007). Introduction to functional data analysis. *Canadian Psychology/Psychologie Canadienne*, *48*(3), 135–155. https://doi.org/10.1037/cp2007014

Locantore, N., Marron, J. S., Simpson, D. G., Tripoli, N., Zhang, J. T., Cohen, K. L., Boente, G., Fraiman, R., Brumback, B., Croux, C., Fan, J., Kneip, A., Marden, J. I., Peña, D., Prieto, J., Ramsay, J. O., Valderrama, M. J., Aguilera, A. M., Locantore, N., … Cohen, K. L. (1999). Robust principal component analysis for functional data. *Test*, *8*(1), 1–73. https://doi.org/10.1007/BF02595862

McKeown G, Sneddon I, Curran W. Gender Differences in the Perceptions of Genuine and Simulated Laughter and Amused Facial Expressions. *Emotion Review.* 2015;7(1):30–38. doi: 10.1177/1754073914544475.

Narayanan, S., Toutios, A., Ramanarayanan, V., Lammert, A., Kim, J., Lee, S., ... & Proctor, M. (2014). Real-time magnetic resonance imaging and electromagnetic articulography database for speech production research (TC). The Journal of the Acoustical Society of America, 136(3), 1307-1311.

Pisanski, K., Cartei, V., McGettigan, C., Raine, J., & Reby, D. (2016). Voice modulation: a window into the origins of human vocal control?. Trends in cognitive sciences, 20(4), 304-318.

R Core Team. (2019). *R: A language and environment for statistical computing.* (3.6.0). R Foundation for Statistical Computing. http://cran.r-project.org/

Ramsay, J. O., Hooker, G., & Spencer, S. (2009). *Functional Data Analysis With R and MATLAB*. Springer Science+Busness Media Inc.

Ramsay, J. O., & Silverman, B. W. (2005). *Functional Data Analysis* (2nd ed.). Springer Science+Busness Media Inc.

Ramsay, J. O., Wickham, H., Graves, S., & Hooker, G. (2017). *fda: Functional Data Analysis* (2.4.7). https://cran.r-project.org/package=fda

Ruch, W., & Ekman, P. (2001). The expressive pattern of laughter. In A. Kaszniak (Ed.) Emotions, qualia, and consciousness (pp. 426-443) World Scientific, Singapore.

Scott, S. K., Lavan, N., Chen, S., & McGettigan, C. (2014). The social life of laughter. Trends in cognitive sciences, 18(12), 618-620.

Waters, S., Kanber, E., Lavan, N., Belyk, M., Carey, D., Cartei, V., ... & McGettigan, C. (2021). Singers show enhanced performance and neural representation of vocal imitation. Philosophical Transactions of the Royal Society B, 376(1840), 20200399.

Wild, B., Rodden, F. A., Grodd, W., & Ruch, W. (2003). Neural correlates of laughter and humour. Brain, 126(10), 2121-2138.

Wiltshire, C. E., Chiew, M., Chesters, J., Healy, M. P., & Watkins, K. E. (2021). Speech Movement Variability in People Who Stutter: A Vocal Tract Magnetic Resonance Imaging Study. Journal of Speech, Language, and Hearing Research, 1-15.