Acoustic analysis of several laughter types in conversational dialogues

Wang, Kexin
Ishi, Carlos
Hayashi, Ryoko

(Citation)
Proc. SpeechProsody 2024:667-671

(Issue Date)
2024

(Resource Type)
conference paper

(Version)
Version of Record

(URL)
https://hdl.handle.net/20.500.14094/0100496620

Speech Prosody 2024
2-5 July 2024, Leiden, The Netherlands

Acoustic analysis of several laughter types in conversational dialogues

Kexin Wang[1,3], *Carlos Ishi*[2,3], *Ryoko Hayashi*[3]

[1]ATR, [2]RIKEN, [3]Kobe University

wkx19970303@hotmail.com, carlos@atr.jp, rhayashi@kobe-u.ac.jp

Abstract

Previous studies suggest the existence of two distinct forms of laughter: mirthful/spontaneous laughter and social/intentional laughter. The current work aims to expand our understanding of the motives behind laughter and its functions in social conversation. About 1000 laughter events from 4 males and 4 females were extracted from multi-speaker conversation data, and the four predominant categories were used for acoustic analysis: mirthful, boosting, smoothing, and softening. Mirthful laughter and boosting laughter exhibit longer duration, higher F0 mean, intensity and HNR, as well as lower H1-A1 than other types, which suggest that laughter produced with positive emotion or attitude tends to be longer, higher and tenser voice quality. On the other hand, smoothing laughter and softening laughter displayed opposite characteristics, which indicates that intentional laughter emitted to smooth the interaction or soften the atmosphere can be acoustically identified to some extent from those with positive emotions. This work provides evidence that laughter with different functions has different acoustic characteristics that help us understand what laughter means in dialogue.

Index Terms: laughter types, laughter functions, acoustic features, voice quality

1. Introduction

Laughter occurs commonly in human interaction and forms an essential part of nonverbal communication. It is not only related to joking and humour, but also plays an important role in social communication such as acting as a sign of politeness or relieving embarrassment [1-4]. Previous studies suggest that laughter can be roughly divided into two different types: mirthful/spontaneous laughter and social/intentional laughter. The former is likely to be elicited by positive moods and expressed toward the dialogue itself, whereas the latter tends to be used to "fill" the conversation and maintain communication with the other [5, 6].

In [7, 8], various types of laughter were categorized into either genuinely mirthful laughter or polite formal laughter. Analysis revealed that, compared to polite laughter, mirthful laughter tends to have longer duration and more calls per bout, as well as shorter bursts, ingressive laugh and chuckle sounds. The perceptual experiment in [6] demonstrated that humans can discriminate social laughter from spontaneous laughter based on auditory or audiovisual presented without any context. Acoustic analysis indicated that spontaneous laughter has not only a longer total duration but also a voiced duration than social ones. However, there was no significant difference in F0 means.

In addition, the neuro-physiological distinctions between these two types of laughter were also proved in brain imaging studies of laughter [9-12]. [9] revealed that the human brain responds significantly differently when listening to authentic amusement laughter and controlled voluntary laughter. The former is characterized by higher activation of bilateral superior temporal gyri, which indicates that listeners process information with more emphasis on the sound itself. In contrast, voluntary laughter was associated with greater activity in the anterior medial prefrontal cortex (amPFC), suggesting that when hearing the posed laughter, listeners try to understand the meaning of it in context or why it happened.

Our recent work reported that spontaneous laughter and intentional laughter tend to have different patterns regarding the number of calls, vowel quality, and voice quality characteristics [13]. For example, spontaneous laughter tends to have more patterns with pressed voice, on the contrary, intentional laughter includes more patterns with vowel /e/ and nasal sounds. However, it was limited to the qualitative analyses of laughter patterns. In the present work, we conducted acoustic analyses to show the differences between two laughter types quantitatively. Furthermore, we analyzed the reasons and functions of laughter in social interaction and compared the acoustic characteristics of four predominant categories.

2. Speech data and annotation of laughter

2.1. Speech data

Most previous laughter studies used induced laughter, for example, funny video clips, rather than spontaneous conversational scenes [14]. To consider the social aspect of laughter, we think it is more appropriate to use a dialogue speech corpus. Thus, in this work, a three-party dialogue database containing several sessions of topic-free conversations with full-channel video and audio was used. The speech data around 12 hours, from four males and four females were used in the analysis. They are all colleagues in the workplace and native speakers of Japanese.

The data recording was conducted in a sound-proof room at the ATR labs, and each conversation session lasts for 15-20 minutes. Three speakers are positioned approximately 1.8 to 2 meters apart from each other. The audio signals are captured using headset microphones (DPA4060). To minimize inter-channel leakages, Wiener filtering is applied among the three channels, utilizing the interfering channels as noise references, so that high signal-to-interference ratios are achieved before acoustic analyses. Additionally, we set a power threshold and do not use acoustic features of frames with low power where the inter-channel leakages are possibly remaining.

2.2. Annotation of laughter

Depending on the positional relationship between speech and laughter, laughter is classified as "stand-alone laughter" (i.e., "ha ha ha") and "speech laughter" (speaking while laughing) [15]. We focus our target on stand-alone laughter. Speech

10.21437/SpeechProsody.2024-135

laughter is also common in the speech data but is not addressed in this work.

According to [14, 16-18], the structure of laughter is hierarchical and usually analyzed at bout, call and segment levels. Segment level is the lowest level which consists of vowels and consonants, or any other changes in production mode usually reflected by the spectrogram components. Call level is syllables consisting of the combination of consonants and vowels or a single vowel. The next higher level is the bout level, which involves at least one call during one exhalation. In many cases, laughter ends with inhalations, but they are not included in the bout level, except for the sounds that are deemed to be critical to the laughter itself, such as ingressive laughter.

We also found "silent segments" which often appear after the regular laughter and before the inhalation, without sound but with the smiling face maintained. Some were also accompanied by head nodding, upper body down, or hand clapping, especially when the speaker is in high tension or struggles to hide laughter. Although the silent segment is not acoustically perceptible, we included them within the bout. An example of annotation of laughter with a 0.85s silent segment is shown in Figure 1.

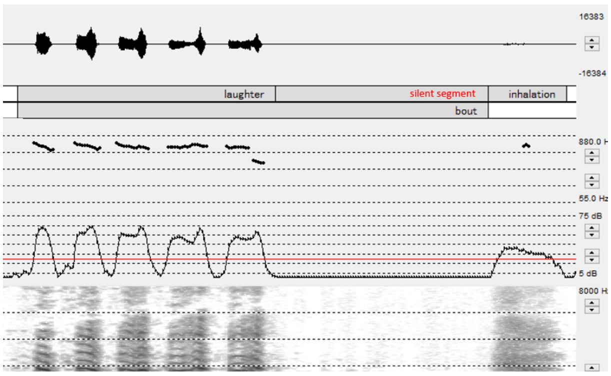

Figure 1: *Example of annotation of laughter with silent segment. (Speech waveform, annotation layers, F0 contour, intensity contour and spectrogram are shown from the top panel.)*

In the present work, the annotation of laughter functions was conducted at the bout level, by segmenting the laughter events based on clear inhalation.

3. Laughter types and laughter functions

The judgement of laughter types (spontaneous or intentional) and laughter functions was conducted by three annotators(research assistants), based on audio and video of the laughter events with 5-second pre- and 6-second post-context. The three annotators have reviewed all conversation data of eight speakers. All three annotators have some degree of familiarity with the speakers, so they have prior knowledge about the personality of the speakers. When the results of the three annotators disagreed, the majority vote was adopted.

3.1. Laughter types

Following previous studies, we categorized laughter into three groups: spontaneous (*sp*), intentional (*in*), and both of them (hereby *sp/in*). Definitions of the type of laughter provided to annotators were: if it seems to you that the speaker laughs in an involuntary manner (*sp*), or in a controlled manner (*in*), or ambiguous (*sp/in*). As a result, out of a total of 857 laughter events, 129 spontaneous laughter, 507 intentional laughter, 152 *sp/in* laughter, and 69 laughter with disagree judgement (i.e., complete disagreement among all annotators) were extracted. Intentional laughter was approximately four times more frequent than spontaneous laughter. *Sp/in* laughter was also a little bit more frequent than spontaneous one.

3.2. Laughter functions

Laughter is expressed in various emotional and intentional states, having different functions in social interaction. Referring to several existing studies devoted to classifying laughter functions [19-22], we used the following labels for annotation and summarized them into 9 categories (Table 1).

Table 1: *Labels for annotation of laughter functions.*

Categories	Labels
Mirthful	funny/joyful/happy
Boosting	boosting/joke/self-mockery/inviting
Smoothing	smoothing/backchannel/contagious/ considerate/conflict/relax
Softening	Softening
Bitter	bitter/deceptive
Self-conscious	self-conscious
Dumbfounded	dumbfounded/derision
Afterglow	Afterglow
Other	awkward/surprised/sympathetic/ anticipatory/etc.

Figure 2 illustrates the result of distributions of laughter functions.

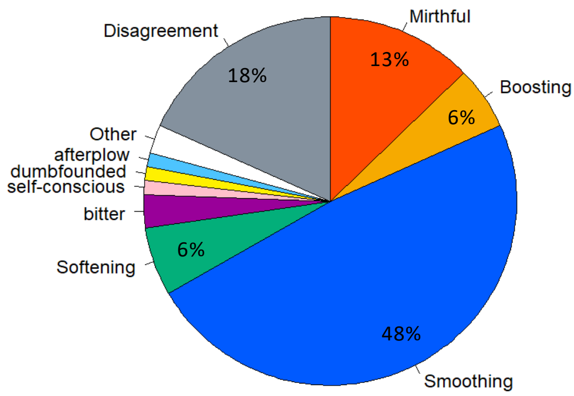

Figure 2: *Distribution of 9 categories of laughter function.*

Excluding the disagreement samples, smoothing laughter occupies the most, approximately half (48%), followed by mirthful laughter (13%), boosting laughter (6%) and softening laughter (6%). The other 5 categories were all less than 3%, therefore they were not used for acoustic analysis. As for the four predominant categories, mirthful laughter is spontaneous, while boosting, smoothing, and softening laughter are intentional. The definitions are shown as follows:

- *Mirthful laughter*: authentic amusement laughter, for example, funny incidents or in a pleasant mood.
- *Boosting laughter*: laughter to liven up the atmosphere of conversation, for example, making it a joke.

- *Smoothing laughter*: laughter to facilitate and smooth conversation, for example, showing comprehension, expressing interest or acting as a backchannel.
- *Softening laughter*: laughter to soften tense, awkward atmosphere or to soften one's statements or arguments.

4. Acoustic Analysis

Acoustic analysis was conducted at the bout level, focused on the difference between spontaneous and intentional laughter, as well as the difference among four predominant categories of laughter functions (mirthful, boosting, smoothing, and softening). Acoustic features including duration, fundamental frequency (F0), intensity, and spectral features *H1-A1* and HNR computed at frame level, in intervals of 10ms. F0 values are converted to semitone units and normalized by subtracting the mean value of each speaker.

4.1. Spontaneous and intentional laughter

Figure. 3 illustrates the differences in acoustic features between the two laughter types based on gender.

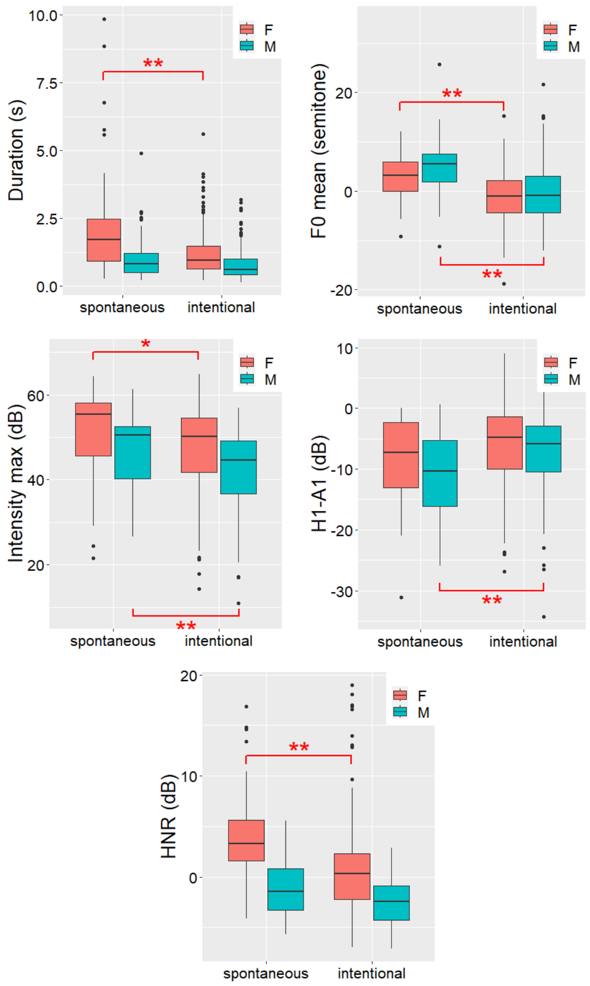

Figure 3: *Mean duration, F0 mean, intensity max, H1-A1, and HNR for spontaneous and intentional laughter at the bout level based on gender.*

The *mean duration* for spontaneous laughter is longer than intentional laughter about 0.96s for females (t = 0.96, p < .01), and 0.25s for males (t = 0.96, *n.s.*). Between spontaneous or intentional laughter, females exhibit longer durations than males. Considering the prevalence of outliers is likely to be due to the varied lengths of silent segment, we compared the duration of silent segment between females and males. Results revealed that the mean duration of silent segment generated by female speakers was 1.09s, which is 0.41s longer than male speakers (t = 8.26, p < .01).

Significant differences are also found between spontaneous and intentional laughter in *F0 mean* and *intensity max* for both female and male speakers. Compared to intentional laughter, the F0 mean of spontaneous laughter is about 4.29 semitones higher for females (t = 4.29, p < .01) and 4.51 semitones higher for males (t = 4.52, p < .01). The intensity max of spontaneous laughter is about 3.47dB higher for females (t = 3.47, p < .01) and 5.44dB higher for males (t = 5.44, p < .01).

H1-A1 (the difference between the amplitudes of the first harmonic and the harmonic closest to the first formant) is a classical measure to discriminate "tense voice" from "lax voice", and to characterize pressed voice segments [23, 24]. Lower values of *H1-A1* indicate high vocal fold tension. Compared to intentional laughter, the *H1-A1* of spontaneous laughter is lower for male speakers (t = 3.81, p < .01), which indicates that spontaneous laughter has a tenser voice quality than intentional one.

HNR (harmonics-to-noise ratio) is a measure of periodicity degree and reflects the effects of the aspiration noise components of breathiness [25, 26]. The *HNR* of intentional laughter is lower than spontaneous one for female speakers (t = 2.96, p < .01), indicating intentional laughter tends to include more elements with high breathiness.

4.2. Mirthful, boosting, smoothing, and softening

We also compared the acoustic features among the four predominant categories of laughter function: mirthful, boosting, smoothing, and softening. Results (Figure 4) revealed that the mean duration and F0 mean of mirthful laughter are longer and higher than smoothing and softening laughter for female speakers. The intensity max of mirthful laughter is higher than smoothing laughter for male speakers, and of boosting laughter higher for female speakers. H1-A1 of boosting laughter is lower than smoothing laughter and softening laughter for female speakers. HNR of smoothing and softening laughter is lower than mirthful and boosting laughter for female speakers.

Based on the above results, a general contrast between mirthful, boosting laughter and smoothing, softening laughter was observed. Regarding duration and F0, the contrast was between mirthful laughter and smoothing, softening laughter. While, as for H1-A1, the contrast was between boosting laughter and smoothing, softening laughter. Moreover, the tendency of intensity and HNR in mirthful and boosting laughter was similar in contrast to smoothing and softening laughter. Although boosting laughter is intentional, as well as smoothing and softening laughter, there are notable differences between them. Analysis results indicate that boosting laughter is more similar to mirthful laughter acoustically.

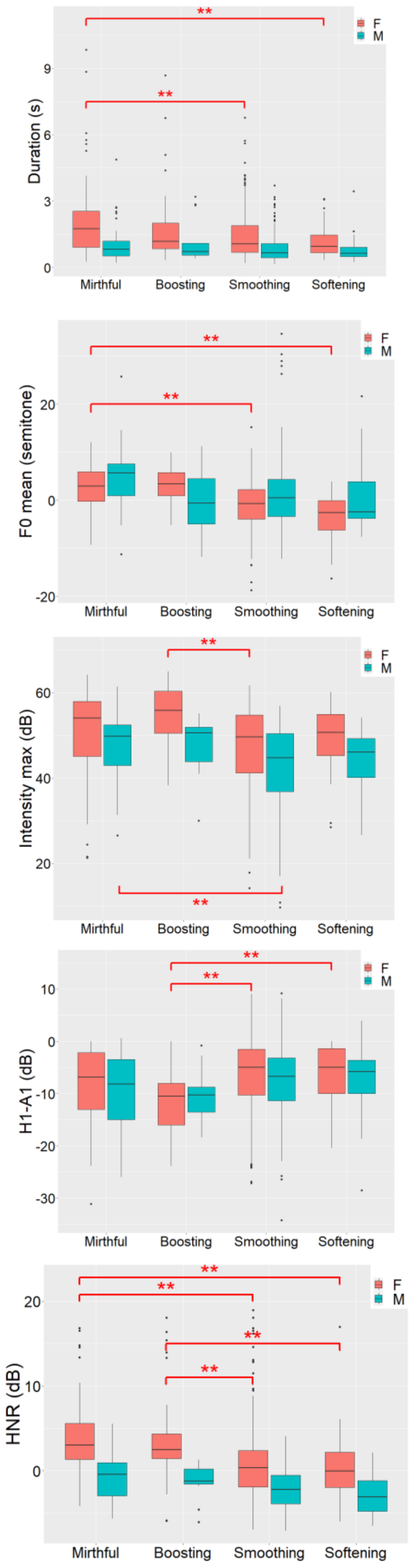

Figure 4: *Mean duration, F0 mean, intensity max, H1-A1, and HNR for mirthful, boosting, smoothing, and softening laughter at the bout level based on gender.*

5. Discussion

The results of the annotation of laughter type and laughter function indicated that most laughter in social interaction is not about pleasantness, but more in an intentional manner, especially to smooth and facilitate conversation, which proves that laughter serves as an important form of nonverbal communication to establish and maintain social bonds with others.

Acoustic analysis revealed that spontaneous laughter tends to be longer, higher and tenser voice quality. This is because when laughing spontaneously, especially in high tension, the air will be squeezed from the lungs with remarkably higher pressure than when produced intentionally. The various differences observed in acoustic features are presumed to be related to physiological distinctions between spontaneous and intentional vocalizations.

Acoustic analysis of four predominant categories of laughter function suggested a contrast between the laughter produced with positive emotion or attitude (mirthful and boosting) and the laughter emitted to facilitate communication or ease tension (smoothing and softening). Intentional boosting laughter has some common features with spontaneous mirthful laughter, such as higher intensity and HNR, and is characterized by a lower H1-A1, indicating that boosting laughter is produced in tenser voice quality due to the effort put into voice. [27] states that voice quality is an important aspect of prosodic information which varies consistently according to the speaker's affect and attitude. This work supported this argument by showing that laughter with different functions was generated in different voice qualities.

6. Conclusions

In this present work, we analyzed the acoustic features of spontaneous and intentional laughter and four predominant categories of laughter function extracted from multi-speaker conversation data. Results revealed that spontaneous laughter has a longer duration, higher F0, intensity and HNR, as well as a lower H1-A1 than intentional one. As for the laughter function, mirthful and boosting laughter produced with positive emotion or attitude tends to be longer, higher, tenser voice quality, and more elements with breathiness. On the other hand, smoothing laughter and softening laughter that emitted to smooth the interaction or soften the atmosphere displayed opposite characteristics.

Regarding gender differences, more significant differences in the laughing styles were observed in females compared to males, which suggests that in laughter production, laughter functions can be distinguished to some extent in females based on acoustic information, while in males conversation contexts or visual modality, such as facial expressions and body movements, might be taken into account. The analysis of other modalities and the aspect of perception will be topics for future investigation. For the acoustic analysis, we also intend to conduct a more detailed analysis of the dynamics features at the call level.

7. Acknowledgements

This work is supported by JSPS KAKENHI Grants (No. JP22H04875, and 21H04348). We also thank Yuka Nakayama, Naomi Uratani and Yumiko Kondo who contributed to the annotations and data analyses.

8. References

[1] R. R. Provine, *Laughter: A Scientific Investigation.* New York: Penguin Books, 2001.

[2] W. Scott, *Ha!: The Science of When We Laugh and Why.* New York: Basic Books, 2014.

[3] J. A. Bachorowski, M. J. Smoski, and M. J. Owren, "The acoustic features of human laughter," *The Journal of the Acoustical Society of America*, vol. 110, no. 3, pp. 1581–1597, Oct. 2001.

[4] J. Vettin, and D. Todt, "Laughter in conversation: Features of occurrence and acoustic structure," *Journal of Nonverbal Behavior*, vol. 28, pp. 93–115 Jun. 2004.

[5] T. Shochi, M. Guerry, H. Suzuki, and M. Kanzaki, "Vocal aspect of social laughter during virtual interaction," *Working Papers em Linguística*, vol. 19, no. 2, pp. 54–77, Dec. 2019.

[6] K. Inoue, D. Lala, and T. Kawahara, "Can a robot laugh with you?: Shared laughter generation for empathetic spoken dialogue," *Frontiers in Robotics and AI*, vol. 9, Sep. 2022.

[7] H. Tanaka, and N. Campbell, "Acoustic features of four types of laughter in natural conversational speech," in *17th ICPhS*, Hong Kong, China, Aug. 2011, pp. 1958–1961.

[8] H. Tanaka, and N. Campbell, "Classification of social laughter in natural conversational speech," *Computer Speech & Language*, vol. 28, no. 1, pp. 314–325, Jan. 2014.

[9] C. McGettigan, E. Walsh, R. Jessop, Z. K. Agnew, D. A. Sauter, J. E. Warren, and S. K. Scott, "Individual differences in laughter perception reveal roles for mentalizing and sensorimotor systems in the evaluation of emotional authenticity," *Cereb Cortex*, vol. 25, no. 1, pp. 246–257, Aug. 2013.

[10] S. K. Scott, N. Lavan, S. Chen, and C. McGettigan, "The social life of laughter," *Trends in Cognitive Sciences*, vol. 18, no. 12, pp. 618–620, May. 2014.

[11] N. Lavan, G. Rankin, N. Lorking, S. K. Scott, and C. McGettigan, "Neural correlates of the affective properties of spontaneous and volitional laughter types," *Neuropsychologia*, vol. 95, pp. 30–39, Jan. 2017.

[12] M. Kosilo, M. Costa, H. E. Nuttall, H. A. Ferreira, S. K. Scott, S. Meneres, J. Pastana, R. Jeronimo, and D. Prata, "Author Correction: The neural basis of authenticity recognition in laughter and crying," *Scientific Reports*, vol. 13, no. 1462, Jan. 2023.

[13] Kexin Wang, Carlos Ishi, Ryoko Hayashi, "Laughter patterns in multi-speaker conversation data," in *20th ICPhS*, Prague, Czech Republic, Aug. 2023, pp.1771-1775.

[14] H. Mori, T. Nagata, and Y. Arimoto, "Conversational and Social Laughter Synthesis with WaveNet," in *20th Interspeech*, Graz, Austria, Sep. 2019, pp. 520–523.

[15] Y. Tian, C. Mazzocconi, J. Ginzburg, "When do we laugh?" in *17th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, Los Angeles, USA, Sep. 2016, pp. 360-369.

[16] M. J. Owren, R. M. Seyfarth, and D. L. Cheney, "The acoustic features of vowel-like grunt calls in chacma baboons (Papio cyncephalus ursinus): Implications for production processes and functions," *the Acoustical Society of America*, vol. 101, pp. 2951–2962, May. 1997.

[17] J. Trouvain, "Segmenting phonetic units in laughter," in *15th ICPhS*, Barcelona, Spain, Aug. 2003, pp. 2793–2796.

[18] K. P. Truong, J. Trouvain, and M. P. Jansen, "Towards an Annotation Scheme for Complex Laughter in Speech Corpora," in *20th Interspeech*, Graz, Austria, Sep. 2019, pp. 529–533.

[19] H. Giles, and G. S. Oxford, "Towards a multidimensional theory of laughter causation and its social implications," *Bulletin of the British Psychological Society*, vol. 23, no. 79, pp. 97–105, Jan. 1970.

[20] C. Mazzocconi, T. Ye, and J. Ginzburg, "What's your laughter doing there? A taxonomy of the pragmatic functions of laughter," *IEEE Transactions on Affective Computing*, vol. 13, no. 3, pp. 1302–1321, 1 Jul-Sep. 2022.

[21] C. T. Ishi, H. Hatano, and H. Ishiguro, "Audiovisual analysis of relations between laughter types and laughter motions," in *8th Speech Prosody*, Boston, USA, May./Jun. 2016, pp. 806–810.

[22] D. P. Szameitat, K. Alter, A. J. Szameitat, C. J. Darwin, D. Wildgruber, S. Dietrich, and A. Sterr, "Differentiation of emotions in laughter at the behavioral level," *Emotion*, vol. 9, no. 3, pp. 397–405, Jun. 2009.

[23] I. Maddieson, and P. Ladefoged, ""Tense" and "lax" in four minority languages of China," *Journal of Phonetic*, vol. 13, no. 4, pp. 433–454, Oct. 1985.

[24] C. T. Ishi, and J. Arai, "Periodicity, spectral and electroglottographic analyses of pressed voice in expressive speech," *Acoustical Science and Technology*, vol. 39, no. 2, pp. 101–108, Mar. 2018.

[25] J. P. Teixeira, C. Oliveira, and C. Lopes, "Vocal acoustic analysis – Jitter, shimmer and HNR parameters," *Procedia Technology*, vol. 9, pp. 1112-1122, Dec. 2013.

[26] I. B. Labuschagne and V. Ciocca, "The perception of breathiness: Acoustic correlates and the influence of methodological factors," *Acoustical Science and Technology*, vol. 37, no. 5, pp. 191-201, Sep. 2016.

[27] N. Campbell, "Changes in voice quality due to social conditions." in *16th ICPhS*, Saarbrücken, Germany, Aug. 2007, pp. 2093-2096.