Laughter patterns in multi-speaker conversation data: comparison between spontaneous laughter and intentional laughter

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

Kobe University[1, 3)], ATR[1,2)], RIKEN[2)], Japan
wkx19970303@hotmail.com, carlos@atr.jp, rhayashi@kobe-u.ac.jp

ABSTRACT

This paper presents the results of an analysis of representative patterns of laughter by 4 males and 4 females, extracted from a corpus of natural conversation speech. In this study, two types of laughter were categorized by a production viewpoint: spontaneous laughter and intentional laughter. The analysis of laughter patterns focused on the number of calls, vowel-quality, and other acoustic characteristics, such as breathy voice and pressed voice. The results showed different trends between the patterns of spontaneous laughter and intentional laughter. A common pattern in spontaneous laughter among all the subjects has been observed, while the patterns of intentional laughter were more individual-dependent and varied, for example, with the usage of nasal sounds and ingressive laughter. However, many of the laughter patterns were found in both spontaneous laughter and intentional laughter, which indicates that other contextual information will also be necessary to understand laughter in dialogue communication.

Keywords: laughter, individuality, spontaneous laughter, intentional laughter, laughter pattern.

1. INTRODUCTION

Laughter is a common behavior throughout life and forms an essential component of nonverbal communication in human social interaction. It is reported that laughter is frequently produced in diverse social situations and with large behavioral variability [1]. For example, [2] classifies laughter into seven types, including humorous laughter, social laughter, ignorance laughter, anxiety laughter, derision laughter, apologetic laughter, and tickling. Moreover, it is reported that there is evident diversity in laughter sex and individual identity [3, 4]. In [3], acoustic analysis was conducted for naturally produced laugh bouts recorded from 97 young adults as they watched funny video clips. The outcomes proved that remarkable variability is found in fundamental frequency characteristics and individuals have their own unique laughter patterns.

On the other hand, the evaluation of synthetic laughter and natural laughter showed that listeners were able to use acoustic characteristics such as laugh duration and F0 to understand the intentions and emotions of the laughing person [5,6]. Furthermore, the contrast between voiced laughter and voiceless laughter has been shown to be an essential difference in laughter variation, with voiced laughter being associated with more positive impressions than voiceless one [7,8].

However, most of the works related to laughter use dialogue data from pairs of speakers, while only a few deal with situations where multiple speakers are present in the environment [9]. We believe that richer phonetic features and laughter patterns exist in multi-speaker conversation data, which is more appropriate for studying the individuality of laughter. Thus, we analyzed laughter in three-way free conversation data in the present work.

2. SPEECH DATA AND ANNOTATION

2.1. Description of the multi-speaker conversation data

A three-party conversation database was used in our analysis. Fig. 1 shows the scenery of the soundproof room where data was collected. The three speakers sat equidistantly around the center table. The database contains several sessions of topic-free conversations of 15 to 20 minutes. Headset microphones and cameras were set up for each subject. Around 13 hours of conversation data by 4 males and 4 females (described as F1-F4 and M1-M4 below) were used in the present laughter analysis. The males are graduate students in their 20s, and the females are research assistants in their 30s to 40s.

Figure 1: The scene of the three-party dialogue data recording.

Depending on the positional relationship between speech and laughter, laughter is classified as "speech laughter" which occurs along with speech utterance, and "stand-alone laughter" which occurs alone without speech [10, 11]. In this work, only stand-alone laughter is targeted.

2.2. Structure of laughter

In [12-15], laughter was analyzed at three primary levels: "segment," "call," and "bout." Fig. 2 shows the hierarchical structure of laughter. At the segmental level, laughter is composed of consonants and vowels. A different segment is recognized when the spectrogram components reflect a clear change in production mode. In actual laughter, it is reported that the first formant is raised and vowels tend to be centralized (schwa) [16]. The call level consists of a combination of consonants and vowels or a single vowel. The call is also called "note" or "syllable." At the bout level, laughter consists of one or more calls, and is typically produced during one exhalation. Although many bouts ended with audible inhalations or exhalations, these sounds were not included in bout-level characterizations unless they were deemed to be critical to the laugh itself [3]. Furthermore, a sequence of multiple bouts is called a laughter episode.

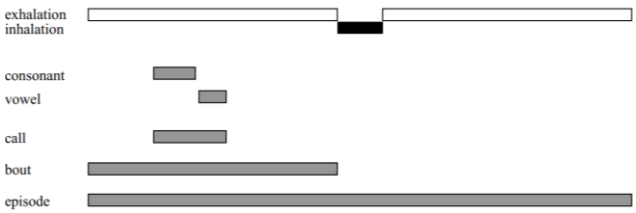

Figure 2: The hierarchical structure of laughter (adapted from [15]).

In this work, we first split the laughter episodes from all speech data, and then split the laughter episodes into bout levels. The following sections 2.3 and 2.4 are both performed at the bout level.

2.3. Spontaneous laughter and intentional laughter

In speech communication, whether the expression is natural or created is a crucial point to be considered [17]. From this production viewpoint, we categorized laughter into two types: spontaneous laughter and intentional laughter.

Spontaneous laughter includes the speaker's involuntary emotion, while intentional laughter is related to the speaker's attitude, intention, and intentionally generated emotion. Unlike spontaneous laughter, when a speaker produces intentional laughter, the listener's reaction might be taken into account for the expression. However, we also recognize the existence of spontaneous laughter with an intentional part, as well as intentional laughter with a spontaneous part. Furthermore, individuality is unavoidably imparted to speech behavior.

The laughter types were annotated by three research assistants, based on audio and video of the laughter events with 5-second pre- and 3-second post-context. Laughter events with two or more matches were used in the analysis. Table 1 shows the number of laughter events per speaker.

Speaker	Spontaneous	Intentional
F1	43	42
F2	16	60
F3	26	26
F4	36	56
M1	22	39
M2	10	9
M3	7	28
M4	21	56

Table 1: The number of laughter events per speaker.

2.4. Annotation of laughter patterns

Taking knowledge from past works on laughter into account, we annotated laughter from the following three viewpoints, as shown in Table 2. Firstly, we divide laughter into four groups by the number of calls: laughter with only 1 call, laughter with 2 or 3 calls, laughter with 4 to 9 calls, and very long laughter with 10 calls or more. Secondly, we describe the vowel-quality of laughter, including Japanese vowels, schwa, and nasal sounds, as well as breathiness. In addition, other prosodic and voice quality characteristics such as pressed voice, falsetto voice, intensity, and some special ways of laughter are also described.

Number of calls		Label 1
1 call		1
2/3 calls		2
4-9 calls		3
10 calls or more		4
Vowel-quality		Label 2
Japanese vowels		a, i, u, e, o
schwa		@
nasal		n
breathiness	pharynx	x
	velar	h
Other voice quality characteristics		Label 3
pressed voice		*
falsetto voice		f
strong (high intensity)		s
weak (low intensity)		w
ingressive laughter		g
whole-voiced (sequence of short calls)		~

Table 2: The annotation of laughter patterns.

The "whole-voiced laugh" item indicates a continuous sequence of calls with short duration, where usually F0 makes small up-down movements between the short calls.

The annotation of laughter patterns is conducted by the first author and two research assistants. The agreement rate between the annotators was not high for laughter patterns (around 50% agreement for two or more annotators). Disagreement mainly occurred in vowel-quality perception (for example, between the vowel /a/ and the schwa vowel). In case of disagreement, the first author's judgment prevailed.

3. ANALYSIS OF THE LAUGHTER PATTERNS

3.1. Variation of laughter call and vowel quality

Table 3 shows the vowel-quality items that appeared in the laughter segments arranged by the laughter type (spontaneous or intentional) and the groups of laughter calls (1 to 4). The phonetic symbols are ordered according to the number of occurrences.

		Vowel-quality
sp (181)	1 (19)	x (14), @ (4), a (2), n (2)
	2 (52)	@ (38), x (23), a (10), n (6), * (4), e (3)
	3 (105)	@ (63), a (45), x (18), * (14), n (3), h (3), f (3)
	4 (5)	@ (4), a (2), * (3)
in (316)	1 (58)	x (47), n (15), @ (10), a (8), e (5)
	2 (147)	@ (67), x (51), n (44), e (14), a (11), h (7), u(5)
	3 (108)	@ (68), x (23), a (18), n (18), e (9)
	4 (3)	@ (3)

Table 3: The distributions of vowel-quality items for different laughter call groups and laughter types. (Number of occurrences are within brackets.)

It can be first observed that laughter call groups 2 and 3 (i.e., the groups of 2/3 calls and 4-9 calls) show the largest occurrences, with group 3 being predominant in spontaneous laughter, and group 2 predominant in intentional laughter.

Regarding vowel-quality, it can be observed that schwa vowels (@) are predominant in all groups except for group 1 (single call), followed by breathiness (x), the vowel /a/, and the nasal sounds /n/. For group 1, the laughter is almost composed of breathy phonation (x). Regarding the call groups 2 and 3, other vowels such as /e/ and /u/, and pressed voice (*) also appear with higher frequency.

We also noticed that vowel-quality could change inside a laughter bout. This change was observed in both spontaneous laughter and intentional laughter, with a tendency for the vowels like /a/, /i/, and /u/ to change into schwa vowels (which can be considered as the default vowel quality during laughter), and for schwa vowels to change into the vowels /a/ (which corresponds to laughter realized by open mouth, mainly in big laughter).

3.2. Relationship between laughter patterns and laughter types

As Table 3 shows, compared to spontaneous laughter, intentional laughter has more vowel-quality variation, especially the vowels /e/ and nasal sounds /n/. On the other hand, spontaneous laughter has more patterns with pressed voice (*).

The typical laughter patterns of each speaker were grouped, and those that occurred twice or more were summarized in Table 4. Fig. 3 shows the waveform, F0 and spectrogram of some representative patterns.

		Typical pattern groups	Other features
F1	sp (43)	1x (5), 2@/2@x (5), 3a (7), 3@/3@x (19)	s (19), * (10)
	in (42)	1x (7), 1@x/1ex/1nx (7), 2@/2@x (4), 3@ (12)	w (6), ~ (9)
F2	sp (16)	2@x (2), 3@/3@x (9)	w (2), g (3)
	in (60)	2@/2@x (14), 2n/2nx (7) 3@/3@x (14), 3e/3ex (6)	s (2), w (5), g (2)
F3	sp (26)	3@ (5), 3a (7)	s (9), g (3), ~ (2)
	in (26)	2a(2), 3@ (4), 3a (3), 3n (3)	s (6), w (2), g (3), ~ (2)
F4	sp (36)	2@ (5), 2a (2), 3@/3@x (11), 3a (3), 3h+g (2)	s (7), w (3), g (3)
	in (56)	1ax (3), 2@/2@x (20), 2n/2nx (12), 3@ (4), 3a (3)	w (9), ~ (3)
M1	sp (22)	3@ (10), 2@ (4), 2@x (3)	f (5), * (7)
	in (39)	1x (5), 1@x (3), 1n/1nx (7) 2@ (10), 3@x(6)	w (5)
M2	sp (10)	2@ (2), 3a+g (4)	s (4), g (4)
	in (9)	2e (3)	
M3	sp (7)	3a+g (2)	g (5)
	in (28)	1x (2), 1@x (2), 2@ (3) 3@+g (8), 3a (2), 4@+g (2)	g (12)
M4	sp (21)	1@x (2), 2@ (3), 2a (2) 2nx (2), 3@+g (3)	s (6), g (6)
	in (56)	1ax/1ex/1nx (6), 2@/2@x (7), 2n/2nx (18), 3@+g (6)	w (10), g (9), ~ (2)

Table 4: Typical pattern groups of spontaneous laughter and intentional laughter. (Number of occurrences are within brackets.)

In spontaneous laughter, call group 3 (4-9 calls) was found to have the largest variety of laughter patterns. 2@, 2@x, 3@, 3@x, 3a are the most typical examples. A common pattern can be observed among the spontaneous laughter of all the subjects, "a ha ha ha", in which the vowel /a/ tends to be centralized and becomes a schwa vowel (such as 3@ example in Fig.

3-3). On the contrary, in intentional laughter, call group 2 (2/3 calls) was found to have the largest variety of laughter patterns. Besides the patterns above, 1x, 2e, 2n, 2nx, and 3n are also commonly found in intentional laughter.

The rightmost column in Table 4 shows the number of occurrences of the prosodic and voice quality characteristics in Label 3, including strong (s), weak (w), ingressive laughter (g), and whole-voiced laugh (~). The strong voice is more predominant in spontaneous laughter, and the weak voice is more predominant in intentional laughter. However, weak spontaneous laughter and strong intentional laughter are also observed. Whole-voiced laugh (Fig. 3-5) is more observed in intentional laughter. Falsetto voice appeared only in M1, but all in spontaneous laughter. Regarding ingressive laughter (Fig. 3-6), no clear trends were observed between spontaneous and intentional laughter.

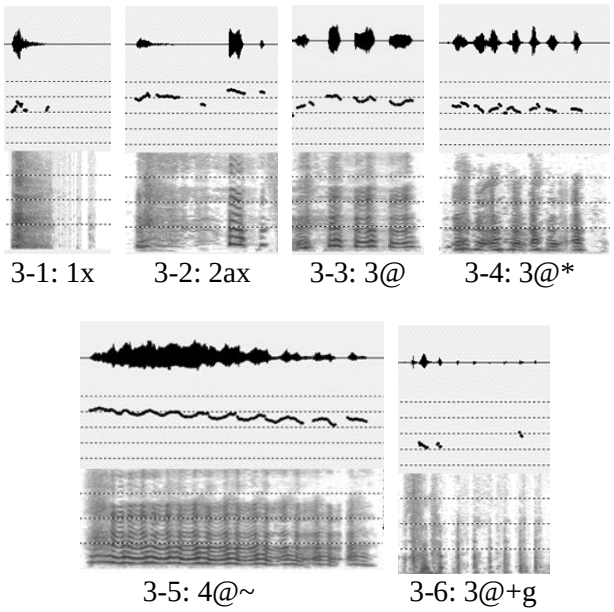

Figure 3: Examples of the representative laughter patterns. Waveform, F0 contour (4 octaves from 55Hz to 880Hz), and spectrogram (8kHz band).

3.3. Individuality of the speakers

The individuality of the speakers regarding laughter patterns can be observed in Table 4. For example, F1 tends to produce pressed voice in spontaneous laughter, and whole-voiced laugh in intentional laughter. M1 tends to use the falsetto voice (f) and pressed voice (*) in spontaneous laughter. F2 and F4 tend to produce patterns with the vowel /e/ and nasal sounds in intentional laughter. In particular, laughter uttered with nasal sounds has a female-like impression and is considered to have a polite effect. Regarding the usage of ingressive laughter (g), F1 and M1 did not use it at all, whereas M3 and M4 made extensive use of a series of ingressive laughter in both spontaneous laughter and intentional laughter. On the other hand, F2, F4, and M2 tend to produce ingressive laughter in spontaneous laughter.

4. CONCLUSION

In the present work, we analyzed the laughter patterns of spontaneous and intentional laughter types in multi-speaker conversation data. The analysis of laughter patterns focused on the number of calls, vowel-quality, and other voice quality characteristics of laughter. The outcomes showed the variation of the laughter depending on the speaker and the social situation.

The laughter with 2 or more calls showed a large number of vowel-quality variations, with the most frequently appearing schwa vowel. In the single-call laughter, breathiness is frequently used. Vowel-quality could change inside one laughter, but usually converges to the schwa or /a/ vowels.

Analysis results also revealed the relationship between laughter types and laughter patterns. Vowel /e/ and nasal sounds /n/ appeared with higher frequency in intentional laughter, whereas pressed voice in spontaneous laughter. A common pattern (“a ha ha ha ha”) was observed in spontaneous laughter among all the subjects, while intentional laughter was more individual-dependent and showed more variations in patterns, such as the usage of nasal sounds, ingressive laughter, and whole-voiced laugh.

Although some trends on the laughter pattern between spontaneous and intentional laughter could be observed, many overlaps in some of the laughter patterns were also observed depending on the speaker. Part of these overlaps are thought to be distinguishable from the visual modality including facial expressions and upper body motions (such visual information was used to classify laughter types, but the analysis of this paper focused on audio-only information). However, others may only be distinguishable from the conversation context. The laughter reason and level of the reason may also be factors to be considered. These are all topics for future investigation. Also, as the present study was limited to qualitative analysis of laughter patterns, our next step is to conduct acoustic analysis to quantitatively show the differences between laughter patterns.

5. ACKNOWLEDGEMENTS

This study was supported by JSPS KAKENHI Grants (No. JP22H04875, 21H04348, and 21K13048). We thank Yuka Nakayama, Naomi Uratani and Yumiko Kondo for contributing to annotations and analyses.

6. REFERENCES

[1] Provine, R. R. 2001. *Laughter: A scientific investigation.* Penguin Books.

[2] Giles, H., Oxford, G. S. 1970. Towards a multidimensional theory of laughter causation and its social implications. *Bulletin of the British Psychological Society*, 23(79), 97–105.

[3] Bachorowski, J.-A., Smoski, M.J., Owren, M.J. 2001. The acoustic features of human laughter. *J. Acoust. Soc. Am.* 111 (3), 1582-1597.

[4] Chafe, W. 2007. *The importance of not being earnest: The feeling behind laughter and humor*. John Benjamins Publishing.

[5] Kipper, S., Todt, D. 2003. The Role of Rhythm and Pitch in the Evaluation of Human Laughter. *Nonverbal Behavior*, 27, 255–272.

[6] Shochi, T., Guerry, M., Suzuki, H., Kanzaki, M. 2019. Vocal aspect of social laughter during virtual interaction. *Working Papers em Linguística*, 19(2), 54-77.

[7] Grammer, K., Eibl-Eibesfeldt, I. 1990. The ritualization of laughter. *Naturlichkeit der Sprache und der Kultur: Acta colloquii*, 192–214.

[8] Owren, M. J., Philipp, M., Vanman, E., Trivedi, N., Schulman, A., Bachorowski, J. A. 2013. Understanding spontaneous human laughter: The role of voicing in inducing positive emotion. *Evolution of Emotional Communication: From Sounds in Nonhuman Mammals to Speech and Music in Man*, 176-190.

[9] Ishi, C., Hatano, H., Hagita, N. 2014. Analysis of laughter events in real science classes by using multiple environment sensor data. *15th Annual Conference of the International Speech Communication Association* Singapore, 1043-1047.

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

[11] Inoue, K., Lala, D., Kawahara, T. 2022. Can a robot laugh with you?: Shared laughter generation for empathetic spoken dialogue. *Frontiers in Robotics and AI*, 9, 155-160.

[12] Owren, M. J., Seyfarth, R. M., Cheney, D. L. 1997. 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*, 101, 2951–2962.

[13] Trouvain, J. 2003. Segmenting phonetic units in laughter. *15th ICPhS* Barcelona, 2793–2796.

[14] Truong, K.P., Trouvain, J., Jansen, M. P. 2019. Towards an Annotation Scheme for Complex Laughter in Speech Corpora. *20th Interspeech* Graz, Austria, 529-533.

[15] Mori, H. Nagata, T. Arimoto, Y. 2019. Conversational and Social Laughter Synthesis with WaveNet. *20th Interspeech* Graz, 520–523.

[16] Szameitat, D. P., Darwin, C. J., Szameitat, A. J., Wildgruber, D., Alter, K. 2011. Formant characteristics of human laughter. *Voice*, 25, 32-37.

[17] Takehara, T. 2019. Emotional communication: facial expressions, voice, body contact, language, cultural differences. In: Uchiyama, I. (eds), Handbook of Psychology of Emotion. Kitaohji Syobo, 147-150.