Understanding Bodily Expressed Laughter Responses to Human and AI-Generated Japanese Manzai scripts

Anonymous submission

Abstract

This study investigates how audiences perceive humor in both Human-created and AI-generated Manzai scripts, a traditional Japanese form of scripted stand-up comedy performed by a pair of comedians. We analyze laughter responses using two complementary modalities: explicit laughter, captured through real-time reaction-button inputs, and implicit laughter, measured via mouth-corner motion tracking. Comparative experiments are conducted to evaluate differences between Human-created and AI-generated Manzai scripts. The results reveal that while AI-generated Manzai scripts can elicit explicit laughter to a certain extent, it is less effective in inducing implicit laughter compared to Human-created script. Implicit laughter is observed only when explicit laughter occurred, suggesting that unconscious emotional reactions may precede conscious recognition of humor. Moreover, Human-created Manzai scripts that are contextually relatable and easy to imagine produced larger and more sustained mouth-corner movements, indicating that empathy and contextual understanding play key roles in embodied humor perception. These findings demonstrate that mouth-corner motion is a reliable indicator of subtle amusement and provide new insights into Bodily Expressed Emotion Understanding (BEEU) for AI-generated humor.

Introduction

Humor is a universal yet culturally nuanced phenomenon that plays a vital role in human communication. It strengthens social bonds, reduces psychological stress, and enhances emotional intelligence. Understanding how humor is perceived and expressed contributes not only to affective computing but also to the study of social cognition and multimodal communication. Previous studies have examined how humans laugh at human-created humorous content, such as humorous dialogue, comedy sketches, and Manzai performances(Platow et al. 2005), (Bryce and Katayama 2009). Researchers have also analyzed when and how audiences laugh while watching human performances based on human-written scripts, using physiological and facial data to capture the timing and type of laughter(Khalid Alnajjar 2022).

However, with the rapid advancement of generative AI, it has become possible to automatically generate humorous content—such as jokes, dialogues, and short stories. As AI-generated humor becomes more widespread and accessible through robots, computer graphics, and chat systems, it is expected to play an increasing role in daily life. Despite this trend, little research has explored how AI-generated humorous content evokes laughter and bodily emotional responses in humans. This study aims to address this emerging research gap.

In this study, we focus on Manzai, a traditional form of Japanese humor. Manzai consists of a dialogue between two performers—a "boke" (funny man) and a "tsukkomi" (straight man)—in which laughter arises not merely from linguistic incongruity, but from the timing, rhythm, and interactive dynamics of their exchange. This dialogic structure provides an ideal framework for generating humor linguistically while also understanding it as a bodily emotional response.

The objective of this study is to investigate how AI-generated Manzai scripts influence human laughter and emotional expression, and to propose an integrated framework that links linguistic creativity with embodied emotional responses. To achieve this, we developed a system that automatically generates Manzai scripts from real-world news articles using partial generative AI. Each script consists of an introduction (tsukami), a main routine (honneta), and a punchline (ochi), composed of multiple "boke components"—pairs of boke and tsukkomi utterances forming a coherent narrative. The AI-generated Manzai script is presented through a chat-based interface with synthesized speech, allowing users to experience the dialogue dynamically as if interacting with virtual comedians. In this study, we compare AI-generated Manzai scripts, produced by our previously proposed AI-based generation system, with Human-created Manzai scripts, written by professional comedians, to clarify their differences in evoking human laughter and emotional responses.

To evaluate human responses, two complementary experiments were conducted.

1. In the **explicit laughter experiment**, users pressed reaction buttons in real time when they found an utterance funny or unfunny, allowing direct annotation of amusement.

2. In the **implicit laughter experiment**, facial motion—particularly mouth-corner movements—was analyzed to detect subtle, unconscious laughter responses. These data were used to compare audience reactions to

AI-generated *Manzai* and professional comedians' performances.

Through these analyses, this study seeks to clarify how AI-generated humor affects human emotion from both linguistic and bodily perspectives. The influence of humor on human emotion involves two major aspects: verbal-humor and non-verbal-humor. Here, verbal-humor refers to linguistic elements such as dialogues and punchlines, whereas non-verbal-humor encompasses gestures, facial expressions, and vocal nuances performed by comedians. As the first step in this research, this paper focuses on the verbal-humor aspect.

Related Work

A considerable amount of research has focused on detecting smiles and laughter from visual, auditory, and physiological cues. Sang et al. (Sang, Cuong, and Van Thieu 2017) developed a convolutional neural network (CNN) that jointly learns smile detection, emotion recognition, and gender classification. Petridis et al. (Petridis and Pantic 2008) proposed a multimodal approach that integrates temporal audio and visual features from conversational data to distinguish laughter from speech. Goumri et al. (Goumri et al. 2023) automatically detected smiles and gaze in children's video calls and compared the performance of traditional feature-based models with end-to-end deep learning methods. Similarly, Bohy et al. (Bohy, El Haddad, and Dutoit 2022) combined audio and video features to classify smiles and laughter using a deep learning framework, and Yang et al. (Yang et al. 2015) applied similar methods to detect laughter and smiles in interactions with older adults. Persona et al. (Persona, Meloni, and Macedo 2023) utilized normalized mouth-landmark coordinates and trained multiple machine-learning classifiers for smile recognition. Huang et al. (Huang et al. 2023) constructed a large-scale smile dataset covering diverse ethnicities and automatically designed an optimized CNN architecture. Kantharaju et al. (Kantharaju, Ringeval, and Besacier 2018) classified laughter types from audiovisual data in dyadic interactions and proposed a method for affect estimation. Scherer et al. (Scherer et al. 2012) automatically detected laughter from audio–video recordings of natural multiparty conversations and compared real-time and offline detection. Melder et al. (Melder et al. 2007) developed an interface called the Affective Multimodal Mirror that detects laughter based on voice and facial expressions, and Jang et al. (Jang, Gunes, and Patras 2019) employed Face-SSD for simultaneous face detection and smile recognition. These studies focus on improving the accuracy of smile and laughter detection. In contrast, we measure mouth-corner motion to analyze which parts of Manzai scripts elicit implicit laughter, with the goal of improving the humor of an automatic Manzai script generation system.

In addition to camera-based approaches, several studies have explored wearable or sensor-based laughter detection. Saito et al. (Saito, Masai, and Sugimoto 2020) embedded photo-reflective sensors in smart glasses to measure smile signals in a contactless manner, and Fukumoto et al. (Fukumoto, Terada, and Tsukamoto 2013) used photo-interrupter sensors to detect smiles and laughter separately. Di Lascio et al. (Di Lascio, Gashi, and Santini 2019) utilized wrist-worn noninvasive wearables to detect laughter from physiological signals and body motion, while Alchieri et al. (Alchieri et al. 2023) investigated electrodermal activity (EDA) asymmetry in wrist data and its effect on laughter recognition performance. Whereas these studies focus on wearable or sensor-based approaches, we measure mouth-corner motion.

Explicit Laughter Response Extraction Method

This study proposes an explicit laughter response extraction method to identify, in real time, which parts of AI-generated Manzai scripts users find amusing. In our previous work, the Manzai Robot system allowed users to enjoy AI-generated comedy through either a dual-robot performance or a web-based chat-style Manzai application(Haraguchi et al. 2021).However, these systems lacked a mechanism to quantitatively capture which utterances elicited laughter during viewing. Instead, user impressions were collected retrospectively through post-viewing questionnaires, leading to a temporal discrepancy between the actual moment of laughter and the reported evaluation.

To address this limitation, we developed a reaction-button-based laughter detection system that extends the existing chat-based Manzai application(show in Figure 1). This system enables users to explicitly indicate their amusement or boredom in real time while watching AI-generated comedy dialogues. As a result, laughter responses can be recorded with high temporal resolution and directly linked to specific utterances in the Manzai script, providing valuable data for analyzing the relationship between dialogue content and perceived humor. The proposed system is implemented as a web-based application that presents Manzai dialogues in a chat-style interface. Two avatars, representing the boke (funny man) and tsukkomi (straight man), alternately perform the dialogue, outputting both the utterance text and the synthesized voice simultaneously. Speech bubbles are rendered using standard style-sheet functionality, and utterances are automatically formatted according to the browser width based on tags embedded in the Manzai script. While viewing the performance, users can indicate their reactions using two types of icons displayed at the bottom of the screen:

- Laughing icon: pressed when the user finds an utterance funny (positive reaction).
- Bored icon: pressed when the user feels the joke falls flat (negative reaction).

These reaction buttons are dynamically generated and synchronized with the progress of the performance, and the interface automatically scrolls as the dialogue continues.

The workflow of the proposed application is as follows:

1. Prepare an AI-generated Manzai script or a Human-created Manzai script.
2. Extract speaker information and utterance text from the script.
3. Generate corresponding speech audio using a text-to-speech engine.

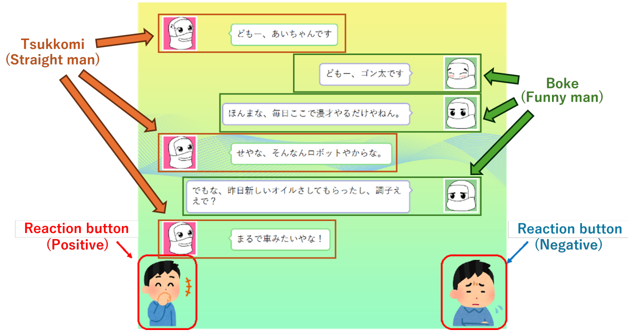

Figure 1: Image of Chat system with reaction buttons

4. The user launches the application and selects the Manzai performance to view.
5. Display the speaker icon and dialogue text synchronized with the synthesized speech.
6. When the user feels amused, they press the laughing icon.
7. When the user feels unamused, they press the bored icon.
8. After each utterance, the system automatically loads the next dialogue segment.
9. Repeat steps 5–8 until the final utterance is completed.

The collected reaction data are stored as explicit laughter responses, which provide direct annotations of users' moment-by-moment amusement. These data will later be integrated with implicit laughter responses (e.g., facial-expression-based laughter detection) to form a comprehensive understanding of how AI-generated or Human-created Manzai scripts evoke human laughter.

Implicit Laughter Response Extraction Method

The explicit laughter response extraction method enables the system to capture, in real time, the moment when a user consciously feels amused during a Manzai performance. However, this approach cannot directly determine whether the user is physically expressing laughter, that is, exhibiting emotional expression through bodily cues. Conventional emotion recognition research commonly analyzes facial expression changes using external cameras. Nevertheless, in a chat-based Manzai application like ours, it is difficult to continuously capture the user's face from the front, making it challenging to obtain stable and accurate facial data.

To overcome this limitation, we propose a laughter detection application designed to objectively extract users' implicit laughter responses. Unlike explicit reactions that depend on button input, this approach captures involuntary, bodily signals of amusement. The proposed method extends beyond chat-based applications and also covers other presentation formats such as robotic performances and live demonstrations. By analyzing users' physical responses, this method automatically estimates which parts of the Manzai evoke laughter, thereby enabling deeper analysis of humor-inducing factors in AI-generated comedy.

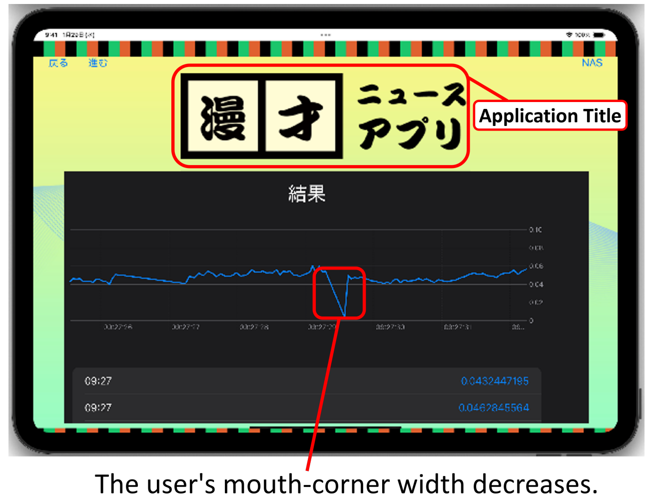

Figure 2: Mouth-corner width decrease indicating an unamused reaction

Most existing facial expression recognition systems classify emotions into predefined categories such as "joy," "sadness," or "anger." However, such categorical classification makes it difficult to directly identify whether a user is "laughing." This is particularly problematic in detecting slight smiles or subtle laughter, which often involve minimal facial changes. Therefore, we focus on mouth-corner motion as a fine-grained quantitative indicator that allows detection of mild laughter, including faint smiles.

For implicit laughter measurement, we employ Apple's Vision 2 framework, which allows real-time tracking of facial landmarks from images captured by the iPad's built-in camera. Among these features, the mouth-corner width—the distance between the left and right mouth corners—is automatically measured. This metric increases notably during laughter and is thus considered an effective cue for detecting subtle amusement responses.

However, facial orientation and posture can cause variations in the measurement. To address this, the application includes an initial 10-second calibration phase, during which the average mouth-corner width is recorded as the user's baseline value. Subsequent measurements are compared to this baseline: a significant increase indicates a laughter onset, whereas a large decrease is recorded as a negative or unamused reaction, as illustrated in Figure 2.

Furthermore, to avoid false detection in cases where the user opens their mouth widely without actually laughing (e.g., speaking or yawning), the system incorporates the inverse of the vertical lip distance as a weighting factor. This correction enhances the detection of upward mouth-corner motion patterns that are characteristic of genuine laughter. To balance real-time performance with computational efficiency, mouth-corner width data are recorded every 0.1 seconds.

Experiments

We conducted two experiments—explicit and implicit—to extract and analyze laughter factors from both Human-created and AI-generated Manzai scripts.

Experimental Conditions

In this study, we analyze how differences in dialogue structure influence laughter generation by comparing Human-created and AI-generated Manzai scripts. Both explicit laughter (collected through real-time reaction buttons) and implicit laughter (measured via mouth-corner motion) were jointly examined to provide a multifaceted understanding of the perception of linguistic humor.

In this experiment, we used three types of Manzai scripts. Two were Human-created Manzai scripts written by professional comedy duos—Milk Boy and Time Machine No. 3—while the other was an AI-generated Manzai script automatically produced by our proposed system. Milk Boy is a well-known Japanese Manzai comedy duo renowned for their logically structured wordplay and precise verbal humor. They won the 2019 M-1 Grand Prix, Japan's most prestigious Manzai competition. Time Machine No. 3 is another popular Japanese Manzai comedy duo, characterized by rapid exchanges and physical expressiveness in their performances.

To minimize bias due to topic differences, all Manzai were unified under the theme of baseball, a familiar and humor-rich subject in Japanese culture. In particular, the professional scripts included routines referencing the Hanshin Tigers, a professional baseball team based in Osaka that is famous for its passionate local fan culture and is often used as a motif in Manzai performances. Table 1 lists the comedians and their respective themes.

The Human-created Manzai scripts were transcribed from recorded performances using our Automatic Manzai Script Transcription System. To isolate the influence of linguistic structure, non-verbal factors such as voice tone, speaking speed, and expressive gestures were kept constant across all performances.

Two laughter detection methods were employed:

1. Explicit laughter response: using the Reaction-Button Laughter Detection Application.
2. Implicit laughter response: using the Mouth-Corner Width Measurement Application.

Twelve participants (university students in their 20s) watched three Manzai performances—by Milk Boy, Time Machine No. 3, and the AI-generated script—in sequence. After each performance, participants answered the following two questions on a five-point Likert scale (5: strongly agree – 1: strongly disagree):

1. Did you find the Manzai funny?
2. Was the Manzai easy to understand?

After viewing all performances, participants were also asked to select which Manzai they found the most entertaining.

The experimental procedure was as follows:

1. Participants watched the Manzai performances in the order of Milk Boy → Time Machine No. 3 → AI-generated script.
2. During viewing, participants provided explicit laughter reactions using the reaction-button application.
3. After each performance, participants completed the post-viewing questionnaire.
4. After viewing all Manzai, participants completed the final comparative evaluation.

Results and Discussion

Explicit Laughter Figures 3–5 show the results of the post-viewing questionnaires. The results indicate that participants generally rated the professional scripts as funnier than the AI-generated script. However, some participants consistently rated all Manzai as ”not funny,” suggesting individual differences in humor perception.

Interestingly, participants rated the AI-generated scripts as easier to understand than those of professional comedians. This finding implies that comprehensibility and funniness are not necessarily correlated in Manzai humor. Moreover, several participants reported that they found the AI-generated Manzai script funnier than the professional ones, suggesting that the AI-generated Manzai script can elicit laughter to a certain degree.

According to the reaction-button data (Table 2), all Manzai achieved over 50% positive laughter responses, confirming that both professional and AI-generated scripts successfully induced explicit laughter among the audience.

In the AI-generated script, one less successful joke included the line:

> ”So, a two-way player means playing guitar with your left hand while making ramen with your right hand, right?”

This joke demonstrates that the AI correctly interpreted the word ”two-way” (nitōryū in Japanese) semantically, but generated unrelated keywords (”guitar” and ”ramen”) that lacked contextual coherence. As a result, participants found it confusing rather than humorous.

In contrast, the professional Manzai scripts often used words that were similar in sound or form to create puns and intentional misinterpretations. For example, a comedian might confuse the word “seal” (the animal) with “seal” (a stamp) or mix up “knight” and “night” to construct a playful misunderstanding. These kinds of phonetic and semantic overlaps are easily recognizable to the audience, allowing them to grasp the intended joke almost instantly. Such deliberate ambiguity between similar-sounding words enables quick mental association and produces laughter through linguistic surprise rather than mere semantic contradiction.

These findings suggest that verbal coherence, phonetic similarity, and contextual association are key linguistic elements that contribute to humor in Manzai. Thus, for AI-generated humor, selecting keywords that are contextually and semantically connected to the preceding dialogue is crucial for eliciting laughter.

Table 1: Comedians and themes used in the experiment.

Comedian	Theme
Milk Boy	Hanshin Tigers (Japanese Professional Baseball Team)
Time Machine No. 3	Shohei Ohtani (Japanese Major League Player)
AI-generated Script	Shohei Ohtani

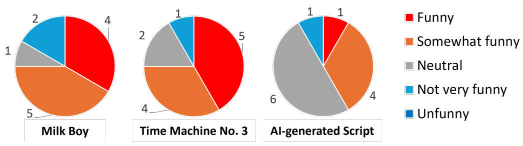

Figure 3: Did you find the Manzai funny?

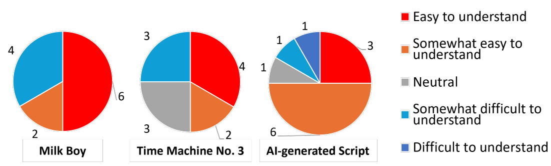

Figure 4: Was the Manzai easy to understand?

Implicit Laughter Figures 6 and 7 present representative results from the mouth-corner width measurement application. For the Human-created scripts, clear fluctuations in mouth-corner width were observed precisely at segments that participants rated as funny, as shown in Figure 6. In contrast, the AI-generated scripts exhibited relatively stable patterns with only minor variations (Figure 7), suggesting that AI-generated Manzai scripts induced fewer implicit laughter responses overall. Nevertheless, closer inspection of the participant data revealed subtle increases in mouth-corner width even during the AI-generated performances. When we examined these local peaks, we found that they frequently occurred in utterances where the content was easy to imagine or relatable to real-life situations (Figure 8 and Figure 9). This indicates that even weak laughter reactions were often associated with cognitive accessibility or emotional resonance. Conversely, segments where the mouth-corner width decreased corresponded to lines that were semantically inconsistent or contextually disconnected from the given topic (Figure 10), as well as to utterances that contained unfamiliar or confusing words that made it difficult for participants to visualize the situation (Figure 11). These results are consistent with the explicit laughter findings: both suggest that when the dialogue lacks contextual relevance or interpretability, humor comprehension declines and laughter—whether explicit or implicit—becomes less likely.

In addition to the explicit laughter responses, this study analyzed mouth-corner width data measured using the Vision 2 framework to examine the tendencies of implicit laughter. In this experiment, participants viewed both Human-created and AI-generated Manzai scripts, and the differences in laughter responses between the two types of content were investigated. The purpose of this analysis was twofold: (1) to determine whether participants' bodily laughter responses corresponded with the moments they explicitly rated as "funny," and (2) to identify differences in laughter reactions between Human and AI-generated Manzai performances.

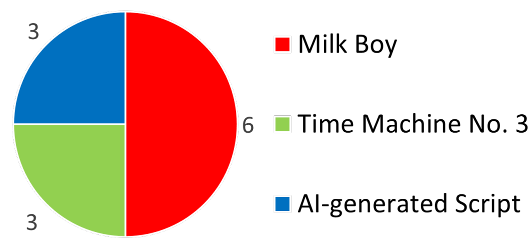

Figure 5: Which of the three Manzai did you find the funniest?

Table 2: Results of the explicit laughter

	Milk Boy	Time Machine No. 3	AI-generated Manzai
Funny	76.7%	61.1%	59.2%
Unfunny	23.3%	38.9%	40.1%

For each participant, mouth-corner width data were continuously recorded at 0.1-second intervals throughout the entire Manzai viewing session. The first 10 seconds of the session were used as a calibration period, during which the average mouth-corner width was defined as each participant's baseline value. Subsequent measurements were compared with this baseline to calculate relative temporal variations. A substantial increase from the baseline was interpreted as a "laughter response," while a notable decrease was regarded as an "unamused reaction." Figures 6 and 7 show examples of temporal changes in mouth-corner width during Manzai performances. Notably, peaks in the implicit laughter data frequently appeared immediately after lines that had been rated as "funny" in the explicit laughter responses, suggesting temporal consistency between the two modalities.

In the case of Human-created Manzai scripts (Figure 6), both the magnitude and duration of mouth-corner increases were larger and more sustained than those observed in AI-generated Manzai scripts. This indicates that professional Manzai performances elicited stronger and more continuous physical expressions of laughter. In contrast, the AI-generated Manzai scripts (Figure 7) exhibited shorter and smaller peaks, suggesting that although the humor was cognitively recognized, it did not evoke strong bodily laughter expressions. These findings imply that, at the current stage,

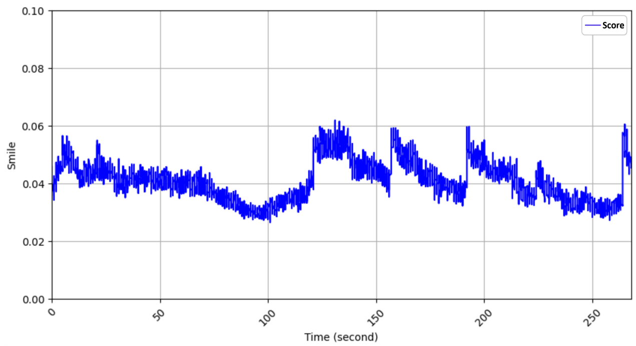

Figure 6: Mouth-corner width measurement results using Human-created Manzai scripts

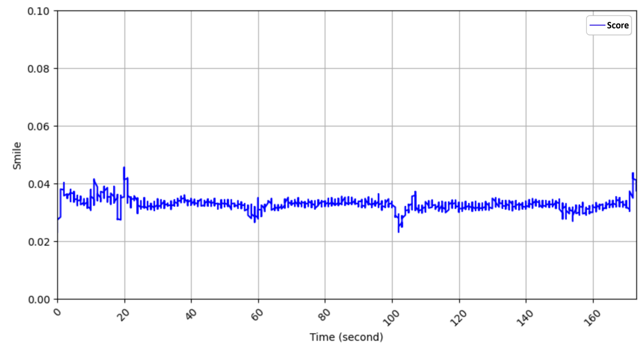

Figure 7: Mouth-corner width measurement results using AI-generated Manzai scripts

AI-generated Manzai scripts has limited ability to induce implicit laughter. However, detailed inspection of the participant data revealed that slight but noticeable increases in mouth-corner width did occur during certain segments of the AI-generated performances. As illustrated in Figures 8 and 9, these upward movements tended to occur when the dialogue was contextually coherent, easily imaginable, or emotionally relatable to the audience. Conversely, decreases in mouth-corner width were frequently observed in scenes such as those shown in Figures 10 and 11, where the "boke" (funny remark) and "tsukkomi" (retort) lines lacked thematic relevance or contained unfamiliar words that made the context difficult to infer. These results suggest that, even in implicit laughter, lexical coherence, contextual predictability, and emotional relatability are critical factors for evoking amusement.

Furthermore, qualitative observation revealed that subtle mouth-corner movements—such as light smiles or suppressed laughter—appeared more frequently in scenes involving linguistic play or verbal wit than in those relying on physical gestures or exaggerated delivery. This finding suggests that implicit laughter is particularly sensitive to linguistic humor, whereas explicit laughter tends to be more influenced by performative and contextual elements. Overall, the implicit laughter results indicate that mouth-corner motion serves as a reliable indicator of subtle amusement. When combined with explicit laughter data, this dual-channel analysis provides a more comprehensive understanding of how audiences perceive humor in both Human and AI-generated Manzai performances.

Tsukkomi	Yeah, well, that's 'cause you're a robot.
Boke	Still, I got an oil change yesterday, so I'm feeling great!
Tsukkomi	What are you, a car?!

Figure 8: AI-generated Manzai script showing increased mouth-corner width.

Total Discussion

To extract both explicit and implicit laughter responses, a comparative experiment was conducted using Human-created and AI-generated Manzai scripts. The analysis of mouth-corner width data revealed that AI-generated Manzai scripts was less effective at eliciting implicit laughter compared to Human-created Manzai scripts. However, a considerable number of explicit laughter responses were still observed in the AI-generated scripts, suggesting that they were capable of evoking cognitive amusement among viewers. Moreover, implicit laughter responses were always accompanied by explicit laughter, whereas the reverse was not observed. This finding implies that laughter may sometimes emerge as an unconscious emotional reaction before individuals consciously evaluate something as "funny."

In the Human-created Manzai scripts, lines such as "Hanshin Tigers fans just want to be ranked above the Giants, and that's enough for them," tended to produce notable increases in mouth-corner width. This result indicates that empathy and contextual comprehensibility play an important role in facilitating bodily expressions of laughter. On the other hand, even in professional Human scripts, there were lines that failed to elicit either explicit or implicit laughter. For example, in the line "When it started raining, we received a huge sheet from the Hanshin Groundskeeping Team to cover the field," the joke relies heavily on accompanying gestures in the live performance to convey the size of the sheet. When presented only as text, however, participants who were not familiar with baseball could not imagine the situation, and thus the humor was lost. This suggests that in Manzai, not only the linguistic content of the script but also performative and embodied elements are essential components for generating laughter.

Furthermore, since implicit laughter was always observed together with explicit laughter, but not vice versa, the proposed implicit laughter extraction method proved to be effective. It serves as a valuable indicator for capturing humor perception through bodily emotional reactions, beyond the level of conscious amusement ratings. These findings highlight the importance of considering not only explicit evaluations but also implicit physiological responses when designing and evaluating future AI-based systems for automatic Manzai script generation that aim to induce laughter.

Boke	So, folks from that team say even if they lose the game, as long as the Giants lose too, they don't get mad.
Tsukkomi	Then that's Hanshin!
Boke	Yep.
Tsukkomi	Hanshin fans are fine as long as they're above the Giants, y'know!
Boke	Uh-huh.
Tsukkomi	To them, a Giants loss counts as a Hanshin win!
Boke	Right.
Tsukkomi	So yeah—must be Hanshin!

Figure 9: Human-created Manzai script showing increased mouth-corner width.

Tsukkomi	On the mound he's throwing heat, at the plate he's launching bombs—peak two-way, man.
Tsukkomi	If the reporters gave him a unanimous vote, that means he's so good nobody can even complain.
Boke	Speaking of "two-way," that's like... playing guitar with your left hand while cooking ramen with your right, right?
Tsukkomi	Yeah, yeah—enjoy the aroma of the ramen while shredding on the guitar...
Tsukkomi	—wait, no! Two-way is about sports!
Tsukkomi	If you're playing guitar while cooking, the noodles'll get soggy!
Boke	But if they get soggy, just dunk 'em back in the soup and they'll bounce back, yeah?
Tsukkomi	They won't!
Tsukkomi	Ramen isn't dodgeball!
Tsukkomi	Now, back to the two-way thing!

Figure 10: AI-generated Manzai script showing decreased mouth-corner width.

Conclusion

This study presented a comprehensive framework for analyzing laughter responses to both Human-created and AI-generated Manzai scripts by integrating explicit and implicit laughter measurements. The explicit laughter responses, collected through a reaction-button interface, captured participants' conscious evaluations of humor, while the implicit laughter responses, measured via mouth-corner motion analysis, provided insights into subconscious bodily expressions of amusement. By combining these two modalities, we successfully demonstrated that humor perception in Manzai involves both cognitive appraisal and embodied emotion.

The experimental results revealed that professional Manzai scripts generated stronger and more sustained laughter responses than AI-generated scripts, particularly in implicit measures. However, AI-generated scripts were still able to elicit laughter in certain contexts, indicating the potential of generative AI to create humorous content. The linguistic analysis further showed that effective humor often relied on contextual word associations, phonetic resemblance, and temporal rhythm—features that current AI models partially capture but do not yet fully master.

Tsukkomi	So, when it started raining just now, the Hanshin grounds crew lent us that tarp they use to cover the field—
Tsukkomi	Thank you very much.
Boke	Thank you.
Tsukkomi	You can never have too many of these—we really appreciate it, seriously.
Boke	Thank you.
Tsukkomi	We managed to get it on somehow, though.

Figure 11: Human-created Manzai script showing decreased mouth-corner width.

These findings contribute to the emerging research field of Bodily Expressed Emotion Understanding (BEEU) by offering a quantitative approach to studying humor as both a linguistic and physiological phenomenon. Moreover, the proposed dual-channel evaluation method provides a foundation for future work on adaptive AI humor systems that can respond to audience emotions in real time.

Future research will extend this framework to multimodal laughter analysis by incorporating voice, gaze, and body gestures, as well as cross-cultural comparisons of humor perception. We also plan to implement the system in robotic or virtual comedian agents capable of autonomously generating and performing Manzai dialogues that adapt to the audience's emotional reactions. Through this line of research, we aim to advance the understanding of human–AI co-creation of humor and to explore how laughter can foster empathy, connection, and well-being in human–machine communication.

Acknowledgments

This work was partially supported by the Premia Project of Konan University, the Research Institute of Konan University, JSPS KAKENHI Grant Numbers 24K03044, and MEXT, Japan.

References

Alchieri, L.; Abdalazim, N.; Alecci, L.; Gashi, S.; Di Lascio, E.; and Santini, S. 2023. On the Impact of Lateralization in Physiological Signals from Wearable Sensors. In *Adjunct Proceedings of the 2022 ACM International Joint Conference on Pervasive and Ubiquitous Computing and the 2022 ACM International Symposium on Wearable Computers*, 472–477. Association for Computing Machinery.

Bohy, H.; El Haddad, K.; and Dutoit, T. 2022. A New Perspective on Smiling and Laughter Detection: Intensity Levels Matter. In *2022 10th International Conference on Affective Computing and Intelligent Interaction (ACII)*, 1–8. IEEE.

Bryce, M.; and Katayama, H. 2009. Performativity of Japanese Laughter. *International Journal of the Humanities*, 6(9): 125–131.

Di Lascio, E.; Gashi, S.; and Santini, S. 2019. Laughter Recognition Using Non-invasive Wearable Devices. In *Proceedings of the 13th EAI International Conference on Pervasive Computing Technologies for Healthcare*, 262–271. Association for Computing Machinery.

Fukumoto, K.; Terada, T.; and Tsukamoto, M. 2013. A smile/laughter recognition mechanism for smile-based life logging. In *Proceedings of the 4th Augmented Human International Conference*, 213–220. Association for Computing Machinery.

Goumri, D.-E.; Janssoone, T.; Becerra-Bonache, L.; and Fourtassi, A. 2023. Automatic Detection of Gaze and Smile in Children's Video Calls. In *Companion Publication of the 25th International Conference on Multimodal Interaction*, 383–388. Association for Computing Machinery.

Haraguchi, K.; Yane, K.; Sato, A.; Aramaki, E.; Miyashiro, I.; and Nadamoto, A. 2021. Chat-type Manzai application: mobile daily comedy presentations based on automatically generated manzai scenarios. In *Proceedings of the 18th International Conference on Advances in Mobile Computing & Multimedia*, 143—-152. Association for Computing Machinery.

Huang, J.; Deng, Y.; Guo, Q.; Xu, Y.; Pan, Q.; and Zhao, Y. 2023. Smile Recognition Based on Comprehensive Dataset Construction and Bayesian Neural Architecture Search. In *Proceedings of the 2023 6th International Conference on Image and Graphics Processing*, 31–35. Association for Computing Machinery.

Jang, Y.; Gunes, H.; and Patras, I. 2019. Registration-free Face-SSD: Single shot analysis of smiles, facial attributes, and affect in the wild. *Computer Vision and Image Understanding*, 182: 17–29.

Kantharaju, R. B.; Ringeval, F.; and Besacier, L. 2018. Automatic Recognition of Affective Laughter in Spontaneous Dyadic Interactions from Audiovisual Signals. In *Proceedings of the 20th ACM International Conference on Multimodal Interaction*, 220–228. Association for Computing Machinery.

Khalid Alnajjar, J. T. J. L. M. K., Mika Hämäläinen. 2022. When to Laugh and How Hard? A Multimodal Approach to Detecting Humor and its Intensity. In *Proceedings of the 29th International Conference on Computational Linguistics*, 6875—-6886. International Committee on Computational Linguistics.

Melder, W. A.; Truong, K. P.; Uyl, M. D.; Van Leeuwen, D. A.; Neerincx, M. A.; Loos, L. R.; and Plum, B. S. 2007. Affective multimodal mirror: sensing and eliciting laughter. In *Proceedings of the International Workshop on Human-Centered Multimedia*, 31–40. Association for Computing Machinery.

Persona, L.; Meloni, F.; and Macedo, A. A. 2023. An accurate real-time method to detect the smile facial expression. In *Proceedings of the 29th Brazilian Symposium on Multimedia and the Web*, 46–55. Association for Computing Machinery.

Petridis, S.; and Pantic, M. 2008. Audiovisual laughter detection based on temporal features. In *Proceedings of the 10th International Conference on Multimodal Interfaces*, 37–44. Association for Computing Machinery.

Platow, M. J.; Haslam, S. A.; Both, A.; Chew, I.; Cuddon, M.; Goharpey, N.; Maurer, J.; Rosini, S.; Tsekouras, A.; and Grace, D. M. 2005. It's not funny if they're laughing: Self-categorization, social influence, and responses to canned laughter. *Journal of Experimental Social Psychology*, 41(5): 542–550.

Saito, C.; Masai, K.; and Sugimoto, M. 2020. Classification of Spontaneous and Posed Smiles by Photo-reflective Sensors Embedded with Smart Eyewear. In *Proceedings of the Fourteenth International Conference on Tangible, Embedded, and Embodied Interaction*, 45–52. Association for Computing Machinery.

Sang, D. V.; Cuong, L. T. B.; and Van Thieu, V. 2017. Multi-task learning for smile detection, emotion recognition and gender classification. In *Proceedings of the 8th International Symposium on Information and Communication Technology*, 340–347. Association for Computing Machinery.

Scherer, S.; Glodek, M.; Schwenker, F.; Campbell, N.; and Palm, G. 2012. Spotting laughter in natural multiparty conversations: A comparison of automatic online and offline approaches using audiovisual data. *ACM Trans. Interact. Intell. Syst.*, 2(1): 1–31.

Yang, F.; Sehili, M. A.; Barras, C.; and Devillers, L. 2015. Smile and Laughter Detection for Elderly People–Robot Interaction. In *Social Robotics: 7th International Conference, ICSR 2015, Paris, France, October 26-30, 2015, Proceedings*, 694–703. Springer-Verlag.