LLT Journal, e-ISSN 2579-9533, p-ISSN 1410-7201, Vol. 27, No. 2, October 2024, pp. 997-1018

LLT Journal: A Journal on Language and Language Teaching
http://e-journal.usd.ac.id/index.php/LLT
Sanata Dharma University, Yogyakarta, Indonesia

PARALINGUISTIC FEATURES IN STUDENTS' STORYTELLING WITHIN MULTIMODAL COMMUNICATIONS

Didik Rinan Sumekto[1*], Hanandyo Dardjito[2], Silvester Goridus Sukur[3], Yunita Susi Andriani[4], Rismalla Ma'im Olam[5], and Dyah Ayu Sulistyorini[6]
[1,2,3,4,5,6]Sarjanawiyata Tamansiswa University, Indonesia
didikrinan@ustjogja.ac.id[1], hanandyo@ustjogja.ac.id[2],
silvestergsukur@ustjogja.ac.id[3], yunitasusiandriani@mmtc.ac.id[4],
rismallaolam@gmail.com[5], and dyah.ayu1@outlook.com[6]
*correspondence: didikrinan@ustjogja.ac.id
https://doi.org/10.24071/llt.v27i2.8701
received 23 May 2024; accepted 6 September 2024

Abstract
This study assessed paralinguistic features in Indonesian university students' contextual interaction during storytelling. Data collection was recorded from respondents' two video-based storytelling performances, while a self-rated questionnaire was distributed to 235 respondents out of 481 undergraduate English education students using simple random sampling. Data analysis used a mixed-methods approach to qualify students' paralinguistic features using the eduistic linguistics annotator (ELAN) and to quantify the paralinguistic features using statistical analyses through the significance of .05. The findings revealed that the ELAN analyzed the contextual interaction among freshmen, sophomores, juniors, and seniors. The students' paralinguistic features corresponded with the lexical and semantic evidence, which approached the function of monosyllabic and bisyllabic words, nonverbal expressions, and interpretations. Bodily gesture quantitatively showed moderate category for 44.7% ($t = 2.434$; $p = .016$), articulation showed attributable category for 54.0% ($t = 3.789$; $p = .000$), facial expression showed moderate category for 61.7% ($t = 2.472$; $p = .014$), and voice loudness showed attributable category for 47.7% ($t = 4.121$; $p = .000$). Herein, positive and significant attribution were shown by these paralinguistic features towards students' contextual interactions in storytelling for 34.9% with the multiple regressions ($F = 7.990$, $R^2 = .349$, and $p < .000$). The paralinguistic features empirically address the multimodal communication modes to improve teaching and learning activities.

Keywords: contextual interaction, multimodal communication, storytelling

Introduction
The function of communication effectively attempts to produce mutual understanding among the users both spoken and nonverbal modes to synergize messages effectively. The latter, nonverbal mode tremendously becomes an option to accomplish paralinguistic features within multimodal communication. Experientially, an effective communication becomes essential to convey ideas between interlocuters and hearers (Kwon et al., 2018), to increase human-to-human

This work is licensed under CC BY-SA.
Creative Commons Attribution-ShareAlike 4.0 International License

communication resolutions, including non-verbal cues which run the consequential function to affect acknowledgment and insight among the language users (Li et al., 2018; Kano, 2018). Some studies on paralinguistic features have contributed to the development of analytically multimodal communications. This, in turn, provides a comprehensive platform for effective communication (Labuschagne, 2017). Paralinguistic features focus on practicing and differentiating students' language achievements (Thacker, Chambers, & Graham, 2018), being well-prepared for language classes, and accommodating them feel self-conscious when delivering messages and ideas (Uştuk & Aydın, 2016). Strengthening effective communication, paralinguistic features respectfully integrate into pedagogical contexts that encourage visual communication modes (Sumekto & Setyawati, 2020), especially when multiple languages are involved (Sumekto et al., 2021).

Paralinguistics' nonverbal communication functionally includes non-lexical cues which encodes, transmits, and decodes from which vocals or sounds, gazes, body languages or postures, rhythms, paces, facial expressions, and gestures are recognized (Cánovas et al., 2020; Gupta, 2016). Nonverbal communication empirically indicates the function of users' conversational flow, emotional expression, and personal and interpersonal attitudes (Gupta, 2016; Pérez-Espinosa et al., 2017), supports visual recognition and language process availabilities, which require particularly conditional settings (Byun, Yoon, & Jung, 2020), and cognitive and non-cognitive loads (Pérez-Espinosa et al., 2017). It also reflects paralinguistic features entrenched in the language signals relating to some similarity and dissimilarity dialects (Sinha, Jain, & Agrawal, 2017) transmitting messages in nonverbal expressions. Speaking competence persists from one subsequent sentence to other substantial sentences to represent the language production (Tooley, Konopka, & Watson, 2018), in which every speaker consequently supports multiple utterances to incorporate an information linkage within classifications (Gosztolya & Tóth, 2018).

Concerned with the intrinsically multimodal human communication that addresses paralinguistic features attribution. Notwithstanding, the features consist of visual communication signals and speech. Bodily gesture drives meanings in both speech presence and absence of co-occurrence. These modalities uniquely utilize movements and configurations to drive meaningful communications (Trujillo et al., 2019) and to indicate the systematic relations between spoken language and gesture. It means that the frequency of distinctly rhetorical designs co-occurrence through structurally- and co-vernacular gesture was measurable to connect the signal parts of equal semantics or communicative functions (Cánovas et al., 2020). Gestures variedly show across one's cultures in relevance with conventions of meaningful symbols, spatial cognition, language, and gestural communication diversity (Kita, 2009; Kwon et al., 2018). Additionally, the bodily gestural expression drives a basic matter of nonverbal communication potentially conveying that genuine human-to-human interaction matters at most.

Some contexts in gestural languages can be more effective than spoken languages communication occurs at a distance since bodily gestural expressions may be pragmatically applied for biometric, identification (personal identification and recognition), and verification (individuals' identity claims confirmation or denial) purposes (Nugrahaningsih & Porta, 2019). Bodily gestures also symbolize conventional hand movements for identifying nonverbal human communication.

These symbolic movements constitute positive or negative social feedback and shall particularly indicate users' emotional expressions (Wood et al., 2018) as naturalistic behaviors (Caridakis et al., 2013). Therefore, users' identifiable and recognizable gestures are dealt with in their cognitive, social, and emotional states (Behera et al., 2020).

In regard to bodily gesture attribution, articulation recognition accuracy supports roles amongst automatic facial expression recognition that deals with speaking subjects. Speech articulation triggers emotion recognition accuracy in speaking (Bursic et al., 2020) and involves individuals' facial variability to convey multiple emotions across different disclosures versus the stability of conveying similar emotions across different disclosures (Slepian & Carr, 2019; Weisbuch et al., 2016), from extracting neutral to peak expressions (Li et al., 2019). These facial expressions readily infer individuals' emotions daily (Barrett et al., 2019). This study also addresses individuals' articulation that controls voice quality in terms of applying for matters upon produced voicing by the vocal folds. This voice generated at the glottis is primarily considered the first element shaping individuals' speech.

Practically, when recognizing interlocuters' voices, they acknowledge the long-lasting things of their accents. They ascertain alterations within short and long vowels, along with consonant fluctuations (Esling et al., 2019) occurring inherently through voice and speech changes (Kaňok & Novotný, 2019). These fluctuations encircle any intonational variations the users express in different identities and are connected with conversational functions (Kozminska, 2018). Another articulation perspective encompasses the vowels output in a phonetic context, which loads nullification and wording. Nullification reduces motion realms and the speaking rate proportion, whereas wording inevitably covers mutual influence of movements within adjoining speech sections (Lierde et al., 2015). Regrettably, wording or articulation disorders may address users' hearing loss, lip injury, motor paralysis, and functional articulation disorders (Aihara et al., 2015).

To address the research gaps in this study, the researchers focused on paralinguistic features in students' contextual interactions in storytelling in terms of addressing bodily gesture, articulation, facial expression, and voice loudness. These features are expressed in the video-based recordings and perceived in the self-rated questionnaire. This study solely engages non-native English students in gaining their contextual interactions through storytelling performance. Previous studies showed that students constructed problems with nonverbal behaviors related to vocal expression, language, and word choices. After drilling three improvised speaking sessions, students drove capabilities of common language use, content, organization, delivery, and confidence (Mortaji, 2018). They also studied how to direct bodily gesture motions, such as their opening palms, hyping index finger, hand orientation, and motion direction (Ishi, Mikata, & Ishiguro, 2020). Behera et al. (2020) pointed out that the assortment of nonverbal behaviors could be connected with modern learning technologies and obtain real-time's affective expressions through computer vision device upon videos captured using a webcam.

On the other hand, gesture may lessen exceptions in accordance with strong roles connected with direction, tasks addressing substance oversight, assignments connected with authentic resemblance, and universally acceptable signs (Wu, 2018). Another influential feature of paralinguistics corresponds with the articulation leading to a strong positive correlation with voice loudness perception.

When the articulation is pronounced loudly, the speech is interpretable to be better (Myers & Finnegan, 2015), and affects facial structures through the auditory, visual, and modality (Bursic et al., 2020). Consequently, facial expression allows interlocuters to analyze relevant social perceptions (Chanes et al., 2018). For example, smiles can be connected with more constructed self-absorbed influence interlocutors' low dynamics, incipience, and high dynamics (Leach & Weick, 2020). Additionally, Park and Stepp (2018) reported that intrinsic pitch was not eligible to be a main effect for vowel identity effects. High dynamics [/i, u/] produces higher vocal tract acoustic impedances relating to a narrower vocal tract than lower vowels.

Hypotheses and research objectives

This study involves hypothesis testing (H_0) with the following paralinguistic features' research questions: (1) How do paralinguistic features support the multimodal communications? (2) Do bodily gesture, articulation, facial expression, and voice loudness both partially and simultaneously attribute a positive and significant attribution towards storytelling's contextual interactions among the Undergraduate university students (Y)? This study aims to analyze the paralinguistic features attribution produced by the Undergraduate students as revealed in their storytelling performances within multimodal communications revelation. Therefore, the interconnected analyses of the independent variables (X_1: Bodily gesture, X_2: Articulation, X_3: Facial expression, and X_4: Voice loudness) towards the dependent variable (Y: Storytelling's contextual interactions) are illustrated in Figure 1.

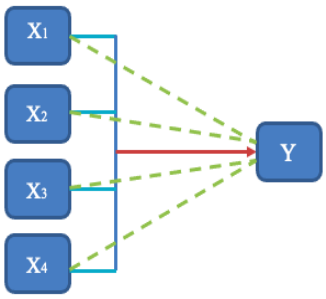


Figure 1. Interconnected analysis: Independent variables toward dependent variable

Method

Participants

This study used the mixed-methods design that constituted paralinguistic features in students' storytelling within multimodal communications to reveal contextual interactions. Four independent variables—bodily gesture, articulation, facial expression, and voice loudness and one dependent variable—students' contextual interactions mainly accomplished this study. This study was set up at a private university in Yogyakarta, Indonesia from which its educational value was affiliated with Ki Hajar Dewantara's teaching. In total, 235 freshmen, sophomores, juniors, and seniors of 481 undergraduate students of the English education program were randomly selected through simple random sampling within each semester. The participant selection process multiply undertook 57 freshmen,

61 sophomores, 53 juniors, 64 seniors, who volunteered to take part anonymously and who were informed that they had the freedom to withdraw themselves at any time. The respondents' category accommodated 76% (n = 179) for females and 24% (n = 56) for males. Meanwhile, the respondents' ages classified 18 to 24 years old (Mage = 21; SD = 4.24) when they participated as the respondents.

Data collection

Data collection used a two-video recording mechanism that showed students' storytelling performance in the speaking class. Prior to recording the participants' storytelling, both male and female participants were intentionally well-informed that their performance would be recorded for data collection in this research. After giving the researchers their permission to record and use their anonymous pictures for study reporting purposes, they started to perform in turns. Two videos entitled, 'Uncle Sam' which was presented by a male student within five minutes and five seconds, whereas 'Fraud and Financial Crimes' was performed by a female student within three minutes and fifty-two seconds. Meanwhile, another data collection accommodated the self-rated questionnaire dissemination towards students' paralinguistics features—bodily gesture, articulation, facial expression, and voice loudness to accomplish storytelling's contextual interactions.

Sample size

Students' responses were measured with a 5-point Likert scale, categorizing from 1 to 5 [1 = not attributable, 2 = less attributable, 3 = moderate, 4 = attributable, 5 = very attributable]. Sample size calculation was used Cohen's formulation [$N = L/f. + U + 1$ to determine the samples. This formulation informed that N = sample size, L = non-centralization parameter, $f.$ = effect size and U = variables] with the significance level [a], power [1-b], effect size [$f.$], and variables [U]. A non-centralization parameter valuation was 15.40, $f. = .15$ [lowest valuation], and 1-b = .90 at .05 significance level (a) with 4 independent variables. Sample size calculation showed N = 15.40 / .15 + 4 + 1; which counted 102.7 + 4 + 1 = 107.7 [decimalizing into 108 for sample size necessities]. Therefore, the sample size determined 235 or 48,85% of freshmen, sophomores, juniors, and seniors in this study to be the respondents.

Validity and reliability

The validity and reliability were assessed by using Cronbach's alpha (α) coefficients, with a threshold of ≥ .70 to be the considerably acceptable value (Griethuijsen et al., 2014). The measurable Cronbach's alpha (α) coefficients showed students' contextual interactions with α = .713, M = 3.97, SD = .371; bodily gesture with α = .867, M = 4.37, SD = .491; articulation with α = .738, M = 3.72, SD = .560; facial expression with α = .849, M = 4.27, SD = .508); and voice loudness with α = .815, M = 3.97, SD = .371.

Goodness of fit test

The Kolmogorov-Smirnov test (K-S Z) was used to withdraw the specifically referenced probability distribution and specifically to compare the equality of one-dimensional probability distribution (p > .05). The results indicated significant relationships between students' contextual interactions (K-S Z = .469; p = .000),

bodily gesture (K-S Z = .415; p = .000), articulation (K-S Z = .387; p = .000), facial expression (K-S Z = .327; p = .000), and voice loudness (K-S Z = .431; p = .000). This suggested that the dependent variable (Y) was aligned with the four independent variables (X_1, X_2, X_3, X_4). Next, the linearity test examined the relationship between independent variables using the F-Calculate (F-Cal.). The results indicated no significant relationship for bodily gesture (F = .843; p = .365), articulation (F = .071; p = .791), facial expression (F = 1.020; p = .371), and voice loudness (F = 1.608; p = .215). Additionally, the homoscedasticity test, conducted using the Glejser test, found no significant influence between independent variables (X) and dependent variable (Y), with p-values greater than .05. Specifically, the results showed no significant relationship for bodily gesture (t-Cal. = -.966; Sig.t = .341; p = .05), articulation (t-Cal. = .259; Sig.t = .797; p = .05), facial expression (t-Cal. = -.885; Sig.t = .383; p = .05), and voice loudness (t-Cal. = -.534; Sig.t = .597; p = .05). The above results showed that there were no multicollinearity implications for the normality, linearity, and homoscedasticity tests.

Data analysis

The eduistic linguistics annotator, so-called ELAN software of version 6.2 from the Max-Planck-Gesellschaft qualitatively analyzed the video recordings relating to the variance of students' storytelling performance. This software empirically identified the annotation mode based on the chronological transcriptions, annotation on the generic mode for Uncle Sam and Fraud and Financial Crimes topic, and media synchronization mode (ELAN, 2021) for male's baritone female's mezzo-soprano voice analyses. Meanwhile, the parametric statistics were used to analyze descriptive (mean and standard deviation), inferential (hypothesizing for X_1, X_2, X_3, X_4) as the independent variables), and multiple regression analyses with its linear regression equation ($Y = a + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4$). Y = dependent variable and a = constant (If X = 0; Y = 0). Hence, the intercepting-Y = 0; b_1, b_2, b_3, b_4 = regression coefficients; and X = independent variables.

Findings and Discussion

Findings

Two videos were analyzed using the eduistic linguistics annotator (ELAN) to assess the paralinguistic features of both male and female students' storytelling performances. The annotations for the two students' videotapes (Figure 2 and Figure 3) were shown chronologically, following a left-to-right model, during the storytelling performance. As stated previously, both student participants took part in the study voluntarily and provided their consent to use their pictures for result reporting purposes. Bodily gestures played a central role in communication for both male and female students, although their gestures had distinct styles. This analysis focused on various components, including references and representations, use of space and movement, and meta-communicative gesticulation. Bodily gestures played a crucial role in real-time multimodal communication. They could be compared to functional paralinguistic features analysis, which encompassed both physical movements and subjective experiences.

In this particular performance, the use of bodily gesture was evident in the way the male student's left fingers and the female student's lips were positioned and moved. These gestures effectively conveyed the story and captivated the audience.

The male student's fingers extended and moved in various ways in front of his chest, enhancing his improvisation and engagement with the narrative on ‘Uncle Sam’.



Figure 2. Male student’s annotation on the generic mode ‘Uncle Sam’ (with photo owner’s permission)

On the other hand, the female student's lips were attractively positioned and delivered different signs as she told a story about ‘Fraud and Financial Crimes’. However, her overall bodily gesture did not contribute significantly to the paralinguistic features, as her hand movements and body mobility were not observable. Nevertheless, her speaking proficiency was impressive and easily understandable.

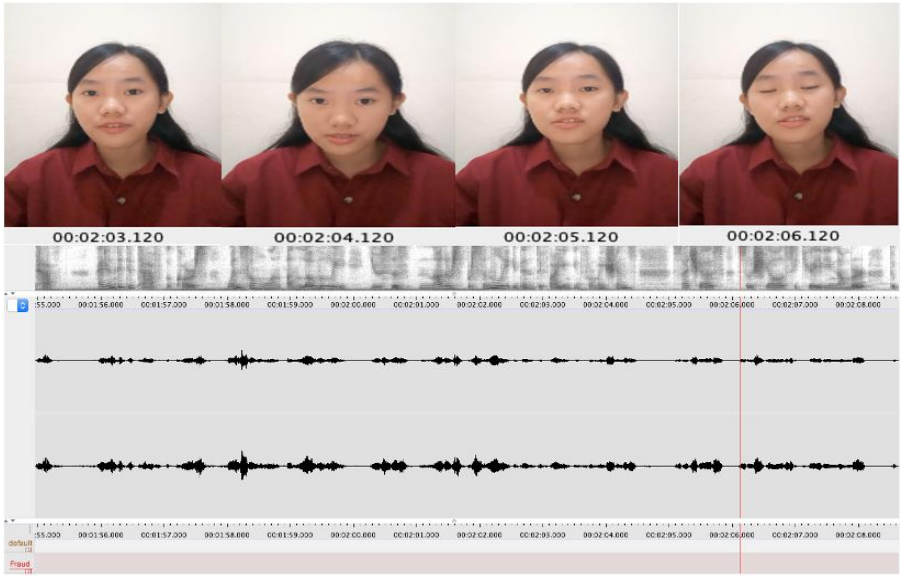


Figure 3. Female student’s annotation on the generic mode ‘Fraud and Financial Crimes’ (with photo owner’s permission)

Articulation assessed students' production upon their speech sounds, including the correctness of their speech utterances and associated motor behaviors. Although most students demonstrated appropriate articulation, some errors were still evident. These errors were primarily related to insufficient intra-oral pressure, resulting in unclear and unacceptable articulation. Figures 4 and 5 highlighted common errors made by both male and female students, such as nasal drag and mid-dorsum palatal stops (a sound that falls between 't' and 'k' in frequency).

Other errors included the incorrect production of velar fricatives (sounds produced through the oral cavity, such as /f, θ, s, ʃ, v, ð, z, ʒ, h/), velopharyngeal fricatives (/s, z, ʃ, ʒ, tʃ, dʒ, tr, dr/), epiglottal stops, and glottal stops (/p…ha/ for /pa/ and /p…ho/ for /po/). Additionally, errors were observed in the vocal tract characteristics, resulting in changes in formant frequencies. Despite these errors, the student's articulation was generally clear and understandable, with only occasional mispronunciations due to limited vocabulary. The student's English fluency was commendable for their basic speaking skills. However, the use of visual aids during storytelling had a moderate impact on the overall comprehension of the student's interactions with their virtual classmates.

Facial expressions could provide valuable insights into a person's emotions and actions. In Figure 4, the male student's nervous facial expression appeared flat, indicating a lack of emotional response. This expression could be attributed to his performance anxiety, physiological stress, motor expression, and potential action tendencies, which could be observed in both positive and negative facial expressions. In this particular case, the male student's facial expression played a dominant role in conveying contextual interactions and serving as a paralinguistic feature in multimodal communication.

Overall, there was a lack of anger, though occasional nervousness could be detected in his voice. The male student kept his eyes fixed on the script in front of him but made an effort to deliver his best storytelling. At times, he gestured with his left hand to enhance his performance. Alternatively, the female student possessed a mezzo-soprano voice that ranged from 'A' below - middle 'C' to two octaves above, with middle 'C' being 'C4' (220-880 Hz). For some cases, the voices could extend as low as 'F' below - middle 'C' (F3, 175 Hz) and as high as high 'C' (C6, 1047 Hz). A clearer voice was strongly associated with greater voice intensity, improved voice transparency, and reduced signs of anxiety during the students' storytelling performance. The variation in voice loudness added color to their speech, influenced by different voice qualities such as modal, whispery, breathy, creaky, harsh, and tense voices.

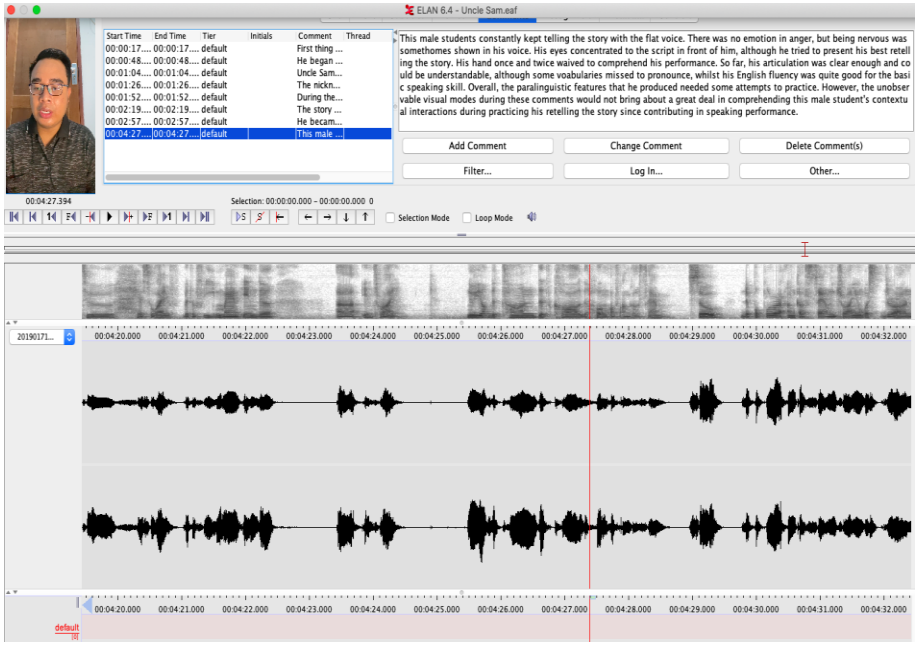


Figure 4. Male student's baritone voice used in media synchronization mode

On the other hand, in Figure 5, the female student confidently expressed herself through her facial expressions while storytelling. Her eye and lip movements clearly demonstrated this confidence. She maintained consistent eye contact with her virtual classmates, conveying a sense of warmth and focused attention. Additionally, her lip movements were noticeable throughout the presentation, providing visual cues to support her communication. Despite her confidence, she did encounter some pronunciation errors with four words, such as 'occur', 'fifth', 'six', and 'racketeering'. This storytelling performance heavily relied on the male students' baritone voices, which ranged from A to F1 and varied in terms of extent, quality, and voice type.

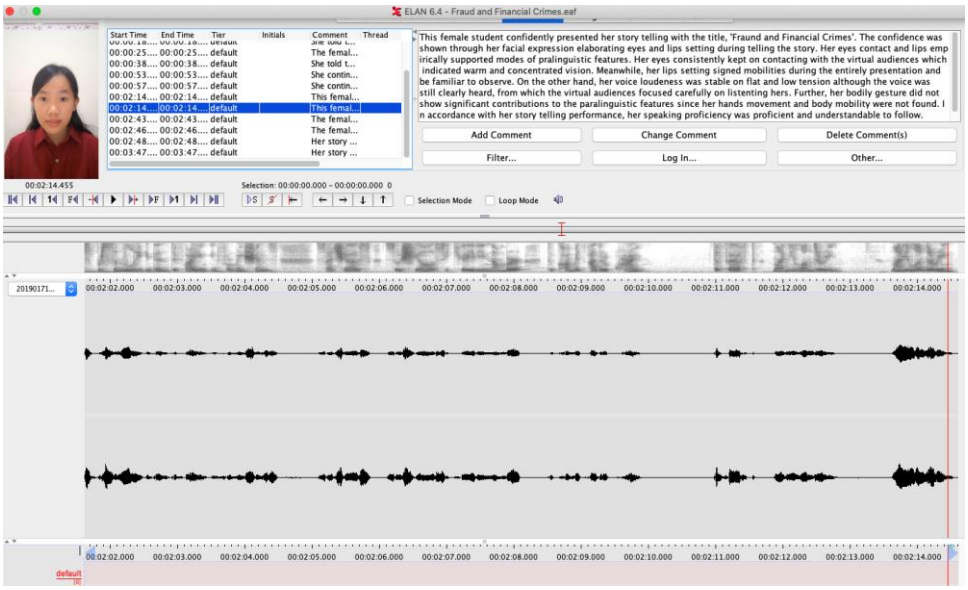


Figure 5. Female student's mezzo-soprano voice used in media synchronization mode

The next analysis focused on the descriptive statistics that interpreted students' contextual interactions in their storytelling session. The students' contextual interactions (Y) were categorized as less attributable (9 or 3.8%), moderate (92 or 39.1%), attributable (104 or 44.3%), and very attributable (30 or 12.8%) of their storytelling session (M = 3.66, SD = .748, N = 235.

Table 1. Frequencies of students' contextual interactions in storytelling

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	2.0 (Less attributable)	9	3.8	3.8	3.8
	3.00 (Moderate)	92	39.1	39.1	43.0
	4.00 (Attributable)	104	44.3	44.3	87.2
	5.00 (Very attributable)	30	12.8	12.8	100.0
	Total	235	100.0	100.0	

Table 1 and Figure 6 confirmed that the highest level of students' contextual interactions in the storytelling session was found to be attributable (104 or 44.3% of the respondents).

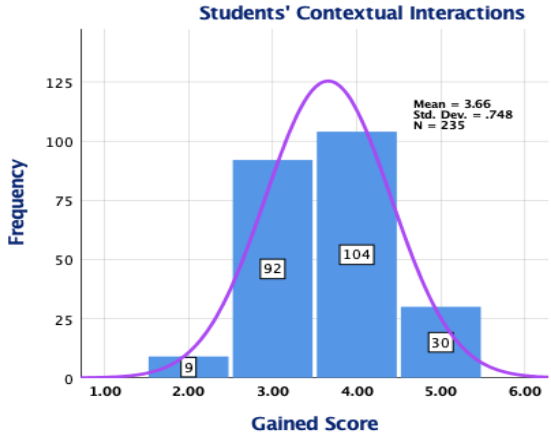


Figure 6. Students' contextual interactions diagram

The bodily gesture (X_1) was analyzed in terms of respondents' perception of mean and standard deviation through the following results: 7 or 3.0% of the respondents found less attributable, 105 or 44.7% of the respondents found moderate, 71 or 30.2% of the respondents found attributable, and 52 or 22.1% of the respondents found very attributable though the bodily gesture feature, where M = 3.71, SD = .842, N = 235.

Table 2. Paralinguistic frequencies of bodily gesture

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	2.0 (Less attributable)	7	3.0	3.0	3.0
	3.00 (Moderate)	105	44.7	44.7	47.7
	4.00 (Attributable)	71	30.2	30.2	77.9
	5.00 (Very attributable)	52	22.1	22.1	100.0
	Total	235	100.0	100.0	

Table 2 and Figure 7 illustrated that the majority of students' bodily gesture in storytelling was attributably assessed as moderate (105 or 44.7% of the respondents).

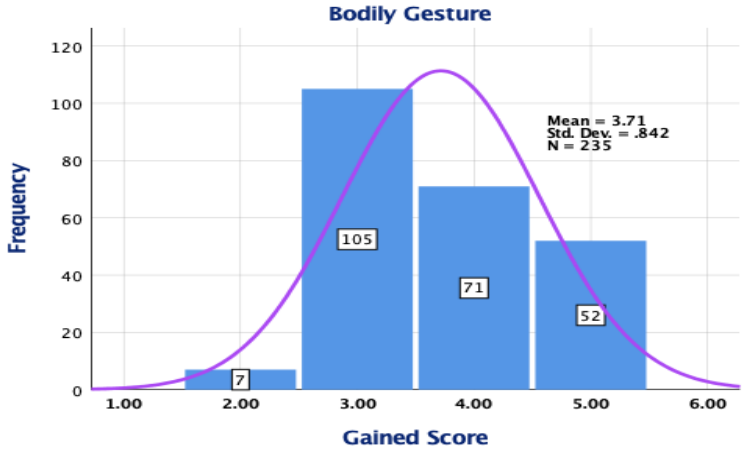


Figure 7. Paralinguistics of bodily gesture diagram

The articulation (X_2) indicated students' perception of mean and standard deviation through the following results: 9 or 3.8% of the respondents perceived it less attribution, 48 or 20.4% of the respondents perceived it moderate attribution, 127 or 54.0% reported attributable attribution, and 51 or 21.7% reported very attributable attribution through the articulation feature, where M = 3.94, SD = .757, N = 235.

Table 3. Paralinguistic frequencies of articulation

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	2.0 (Less attributable)	9	3.8	3.8	3.8
	3.00 (Moderate)	48	20.4	20.4	24.3
	4.00 (Attributable)	127	54.0	54.0	78.3
	5.00 (Very attributable)	51	21.7	21.7	100.0
	Total	235	100.0	100.0	

Table 3 and Figure 8 showed that the majority of students' articulation in storytelling was assessed as attributable (127 or 54.0% of the respondents).

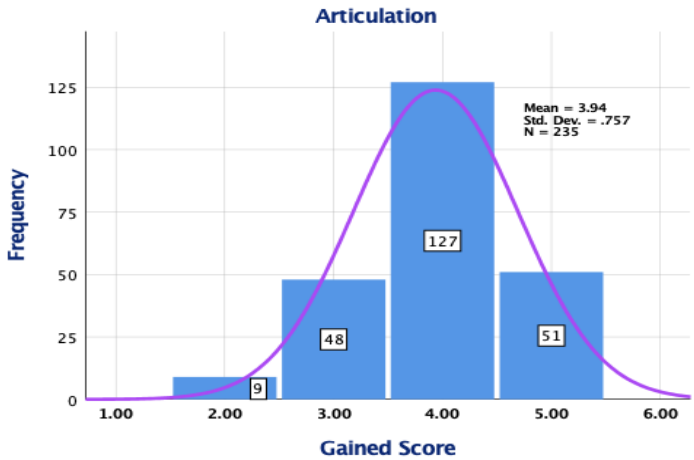


Figure 8. Paralinguistics of articulation diagram

Furthermore, the facial expression (X_3) indicated that students' perception of mean and standard deviation categorized in the following categories: less attributable for 14 or 6.0% of the respondents, moderate for 145 or 61.7% of the respondents, attributable for 37 or 15.7% of the respondents, and very attributable for 39 or 16.6% of the respondents through the facial expression feature, where M = 3.43, SD = .836, N = 235.

Table 4. Paralinguistic frequencies of facial expression

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	2.0 (Less attributable)	14	6.0	6.0	6.0
	3.00 (Moderate)	145	61.7	61.7	67.7
	4.00 (Attributable)	37	15.7	15.7	83.4
	5.00 (Very attributable)	39	16.6	16.6	100.0
	Total	235	100.0	100.0	

Table 4 and Figure 9 showed that the majority of students' facial expressions were attributably assessed as moderate (145 or 61.7% of the respondents).

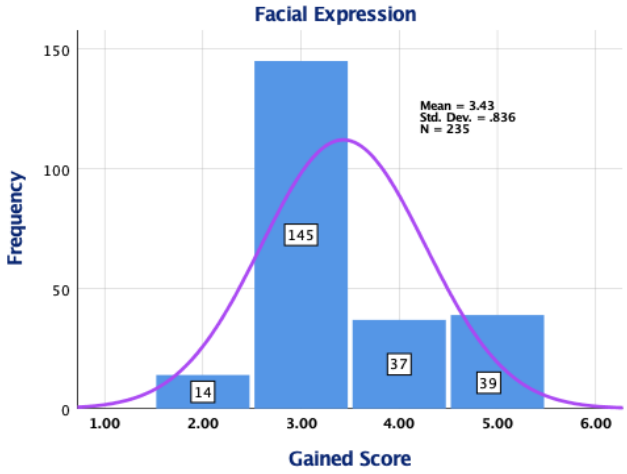


Figure 9. Paralinguistics of facial expression diagram

Finally, the voice loudness (X_4) indicated that students' perception of mean and standard deviation categorized: less attributable for 7 or 3.0% of the respondents, moderate for 62 or 26.4% of the respondents, attributable for 112 or 47.7% of respondents, and very attributable for 54 or 23.0% of the respondents through this voice loudness feature, where M = 3.91, SD = .779, N = 235.

Table 5. Paralinguistic frequencies of voice loudness

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	2.0 (Less attributable)	7	3.0	3.0	3.0
	3.00 (Moderate)	62	26.4	26.4	29.4
	4.00 (Attributable)	112	47.7	47.7	77.0
	5.00 (Very attributable)	54	23.0	23.0	100.0
	Total	235	100.0	100.0	

Table 5 and Figure 10 showed that the majority of students' voice was attributable (112 or 47.7% of the respondents).

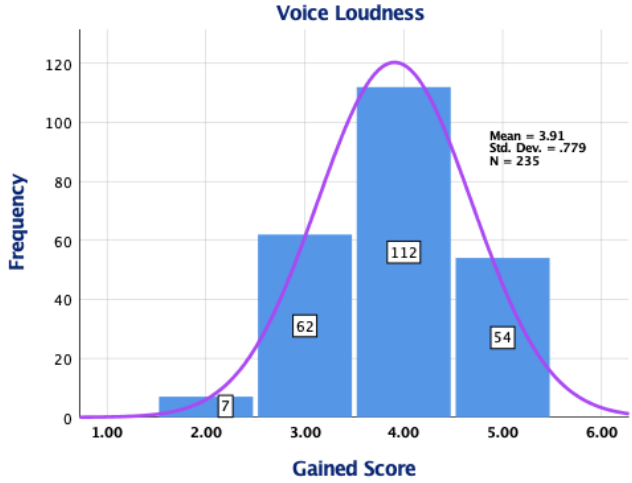


Figure 10. Paralinguistics of voice loudness diagram

Additionally, the standard multiple regression examined the connection of storytelling's contextual interactions with freshmen, sophomores, juniors, and seniors within multimodal communications and empirical paralinguistic features. The regression analysis revealed that bodily gesture ($F = 5.924$; $p = .016$), articulation ($F = 14.355$; $p = .000$), facial expression ($F = 6.111$; $p = .014$), and voice loudness ($F = 16.984$; $p = .000$) derived a significant role in their attributions. The multiple regression analysis (Table 6 and Figure 11) showed that the step-wise method used to reveal the linear regression equation ($Y = a + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4$) with its equation values, $Y = 1.670 + .071X_1 + .159X_2 + .098X_3 + .196X_4$. This empirical relevance was measured by the determinant coefficients, where the multiple $R = .349$ and $p < .05$. This meant that bodily gesture (X_1), articulation (X_2), facial expression (X_3), and voice loudness (X_4) contributed positively and significantly to students' contextual interactions (Y). These independent variables contributed 34.9% of the variation in contextual interactions, while the remaining calculation of 65.1% was still potentially attributable to other paralinguistic features in further research.

Table 6. Multiple regression analysis

Paralinguistics Features	**β**	**r^2**	**t**	***p* (Sig.)**
Bodily gesture (X_1)	.017	1.239	2.434	.016
Articulation (X_2)	.159	2.448	3.789	.000
Facial expression (X_3)	.098	1.756	2.472	.014
Voice loudness (X_4)	.196	3.207	4.121	.000
Constant = 1.670			Alpha (α) = .05	
Multiple-R = .349			R^2 (Square) = .122	
F = 7.990				*p* = .000

In this section, the regression's standardized residuals (Figure 11) consistently followed the diagonal line, indicating that the residual values had a normal distribution. Therefore, we could conclude that the assumption of normality for the residual values in the standard multiple regression was valid. On the other hand, the scatterplot of students' contextual interactions (Figure 11) revealed that there was no evidence of heteroscedasticity in the regression model. The scatterplot showed various patterns above and below the null (0) on the Y-axis, indicating the absence of heteroscedasticity. However, it was worth noting that the straight line sloped downward from left to right, indicating that as the independent variables (X_1, X_2, X_3, X_4) increased, whereas the dependent variable (Y) decreased proportionately, as resulting in a negative correlation.

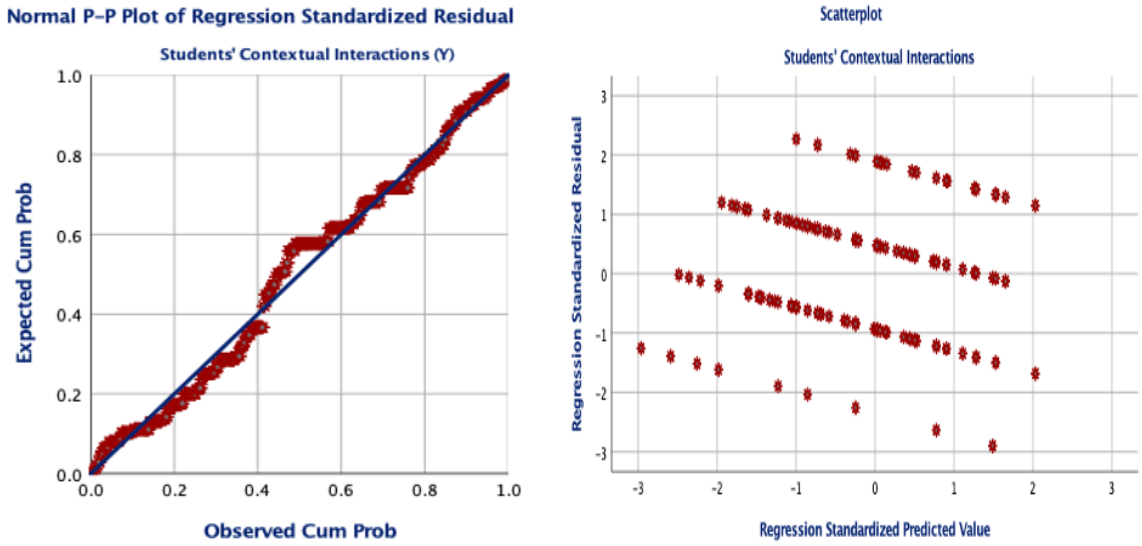


Figure 11. The normal P-P Plot of regression standardized residual and scatterplot of students' contextual interactions

Discussion

When it comes to paralinguistic features, the complex nature of speaking proficiency has a significant impact on factors such as individual ethnicity, culture, and existing competence levels. Speaking proficiency is closely tied to both team and individual processes of speaking competence and recognition (Li et al., 2018). The freshmen, sophomores, juniors, and seniors rely on thinking-for-speaking patterns, including the conceptualization and construction of manner and motion paths, to effectively express themselves verbally. Meanwhile, co-speech gesture enrols an important position in non-native English students' assistance. It challenges to conceal difficulties in a way that can still be understandable (Wessel-Tolvig & Paggio, 2016). Another paralinguistic feature aspect conveys the identification and decoding of verbal information among the students, which greatly aids in comprehensive pronunciation (Ali, Segaran, & Hoe, 2015). Additionally, nonverbal vocalizations, such as co-laughter and aggressive confrontations are prevalent across different cultures (Šebesta et al., 2019).

The students intentionally afford the paralinguistic features in the storytelling within multimodal communications to support their contextual interactions. The interactions demonstrate their ability to understand and categorize paralinguistic

symbols, both in context and out of common context. Paralinguistic features—bodily gesture, facial expression, and voice loudness enrols an important position in conveying meaning alongside the linguistic elements, like morphology and syntax constructions. The students realize that these paralinguistic features experientially enhance their speech's contextual interactions within the expressive rhythm. In other words, they experience with speech and paralinguistic features which blend in their utterances. Overall, these contextual interactions shape students' language accommodation and encompass physical, psychological, and social aspects. By examining students' storytelling performance, this study attempts to understand how their fluency, pronunciation, contextual understanding, vocabulary, and grammar produce together through the selected topics like Uncle Sam and Fraud and Financial Crimes.

Students' multimodal communications intentionally support their contextual interactions, portray both cognitive and affective competencies to figure out conditional categorization and decontextualization of verifiably linguistic symbols. As the symbols of acknowledged paralinguistic features, these empirically govern the morphological, syntactical, implicit, and explicit meanings in the hierarchical structure for comprehension. Students conditionally mobilize these paralinguistic features more than just speaking contextually, but expressing themselves rhythmically. In short, students also possibly create utterances that combine both speech and paralinguistic features. In this experience, contextual interactions conceptualize language use, broadly defining physical, psychological, and social matters from which spoken expressions are accommodated. To derive these contextual interactions, students' storytelling performance produces a particular situation that concretizes any venue where the conversation takes place, while language experience indicates their fluency, pronunciation, contextual content, vocabulary, and grammar.

Apart from working with standard multiple regression analyses, bodily gesture partially shows a significant attribution ($t = 2.434$; $p = .016$), while its regression equation ($Y = 3.140 + .017X_1$) is graded into the moderate category (105 or 44.7% of the respondents). Meanwhile, null hypothesis (H_0) assumption confirms that H_0 = there is no positive and significant attribution influencing to bodily gesture (X_1) toward students' contextual interactions (Y). On the other hand, Ha = there is a positive and significant attribution influencing to bodily gesture (X_1) toward students' contextual interactions (Y). Therefore, H_0 is not acceptable if the p-value is less than .05 ($p < .05$), whereas bodily gesture coefficients show (R) = .016 and $p < .05$. Based on the result, H_0 is not acceptable, while alternative hypothesis (Ha) is acceptable. Herein, bodily gesture shows a positive and significant attribution toward students' contextual interactions within its coefficients (R) = .016. Bodily gesture production tends to be significant for planning discourses and propositions, as well as reaches users' self-oriented cognitive functions (Lin, 2020). Bodily gesture indicates communication to gain social interaction skills and shows behavioral uniqueness in motor, cognition, language, social connections (Kwon et al., 2018), collective actuation, acceleration, dominion, fluency, infinite boundary, and natural interaction parameters as bodily gesture expressivity (Caridakis et al., 2013), as well as produced bodily natural movements (Pichel, 2016). The productive and incorporated bodily gesture's multimodal communications utilize communication and maximize integrated

informativity of the multimodal signals (Cánovas et al., 2020). The direction, order, familiar objects, and performable tasks may be easier for the English users to understand the adopted gestures (Wu, 2018), particularly when cognitive skills contribute to modulating gesture across ages (So & Wong, 2017). So far, the representational bodily gesture affects cognitive skills in activation, manipulation, packaging, and spatial-motoric information for thinking and speaking exploration from generating practical actions (Kita, Alibali, & Chu, 2017). Hence, students' bodily gesture potentially stimulates numerous interpretations, knowledge, and meaningful comprehension (Sumekto et al., 2021).

Meanwhile, articulation shows a significant attribution ($t = 3.789$; $p = .000$), while its regression equation ($Y = 2.723 + .159X_2$) is graded into the attributable category (127 or 54.0% of the respondents). Meanwhile, null hypothesis (H_0) assumption confirms that H_0 = there is no positive and significant attribution influencing to articulation (X_2) toward students' contextual interactions (Y). On the other hand, Ha = there is a positive and significant attribution influencing articulation toward students' contextual interactions (Y). Therefore, H_0 is not acceptable if the p-value is less than .05 ($p < .05$), whereas articulation coefficients showed (R) = .000 and $p < .05$. Based on the result, H_0 is not acceptable, while alternative hypothesis (Ha) is acceptable. Herein, articulation shows a positive and significant attribution toward students' contextual interactions within its coefficients (R) = .000. Articulation is understandable as the acoustic measure that investigates pitch and resonance, whereas the auditory-perceptual measure examines tempo and stress. So, voice quality and intonation become part of auditory perceptual and acoustic measures (Leung et al., 2018) in articulation as well. The broad-minded maturation of the auditory system is eligible to affect speech perception among students, to resolve spectral and temporal details in the indispensable speech sounds (Cabrera, Lorenzi, & Bertoncini, 2015). What is found about articulation in this study relies on provisional reflection and linguistic communication guidance, visual, and coexisting communication among students' articulation (Sumekto et al., 2021). Hence, future studies might deliberate the vocal fold efficiency, generalize strategies to speaking issues, and distinguish the loudness conditionally (Myers & Finnegan, 2015).

Furthermore, facial expression partially shows a significant attribution ($t = 2.472$; $p = .014$), while its regression equation ($Y = 3.169 + .098X_3$) is graded into the moderate category (145 or 61.7% of the respondents). Meanwhile, null hypothesis (H_0) assumption confirms that H_0 = there is no positive and significant attribution influencing to facial expression (X_3) toward students' contextual interactions (Y). On the other hand, Ha = there is a positive and significant attribution influencing facial expression toward students' contextual interactions (Y). Therefore, H_0 is not acceptable if the p-value is less than .05 ($p < .05$), whereas facial expression coefficients show (R) = .014 and $p < .05$. Based on the results, H_0 is not acceptable, while alternative hypothesis (Ha) is acceptable. Herein, facial expression shows a positive and significant attribution toward students' contextual interactions within its coefficients (R) = .014. Facial expression empirically derives social interpretations of stereotypically non-affiliative emotions. For example, a frowning face can be evaluated more positively in an angry scenario when someone is stuck in traffic (Chanes et al., 2018). Facial expression reflects abilities to relive emotions from experience and to comprehend future thinking (Haj et al., 2018). So,

it is affected across all portraits with reduced low power and increased high power (Leach & Weick, 2020). In understanding oral and non-verbal functions, students' facial expression can show postures and eyes contact as evidence-based expressions (Sumekto et al., 2021) and are increasingly recognized as comprehensible paralinguistics that allow students in resuming better explicit and implicit meanings, particularly in speaking through video-based performance (Labuschagne, 2017).

Lastly, voice loudness partially shows a significant attribution ($t = 4.121$; $p = .000$), while its regression equation ($Y = 2.682 + .196X_4$) is graded into the attributable category (112 or 47.7% of the respondents). Meanwhile, null hypothesis (H_0) assumption confirms that H_0 = there is no positive and significant attribution influencing to voice loudness (X_4) toward students' contextual interactions (Y). On the other hand, Ha = there is a positive and significant attribution influencing voice loudness toward students' contextual interactions (Y). Therefore, H_0 is not acceptable if the p-value is less than .05 ($p < .05$), whereas voice loudness coefficients show (R) = .000 and $p < .05$. Based the results, H_0 is not acceptable, while alternative hypothesis (Ha) is acceptable. Herein, voice loudness shows a positive and significant attribution toward students' contextual interactions within its coefficients (R) = .000. Voice loudness modality may perform closely equivalent to the arousal classification, although low arousal is better detected. This seems flimsily better fitted for relative capacity to unite recognition (Caridakis et al., 2013). To support students' voice loudness, the airflow vibrato forms a strong relationship with pitch. A wider airflow vibrato extends higher pitch, particularly females tend to produce greater vibrato rather than males (Nandamudi & Scherer, 2019). The substantial frequency of vocal folds opening and closing slightly increases the opportunity for particles to generate larynxes (Asadi et al., 2019), while utterances are primarily interpretable as students' characteristics to produce voice loudness either (Šebesta et al., 2019).

Paralinguistic features in students' storytelling within multimodal communications accordingly diminishes foreign language anxiety in terms of communication apprehension, negative evaluation fear, panic prevention, confusedness, and bodily advancement upon any anxieties. Further, paralinguistic features considerably increase feelings of comfortability, self-concept of producing proficient languages, parole proficiency, linguistic skills, and impetus for class preparation (Uştuk & Aydın, 2016). They notably impact hyperbolized styles, such as lack of controls, disarranged interlocutors' etiquette as opposite controls sprightly (Nilsen et al., 2016). In this study, students' spoken expressions are primarily influenced by their distinct knowledge alterations involving speech fluency cues through the referential expectations formation for words recencies, as if they still hesitate with disfluency dynamics in the unfolding utterances (Thacker, Chambers, & Graham, 2018). Its attribution corresponds with full lexical and semantic information which can lead to monosyllabic or bisyllabic words, as well as influence voice recognition on the phonological familiarity for voice recognition cruciality (Zarate et al., 2015).

Conclusion

Paralinguistic features assess the storytelling performance of freshmen, sophomores, juniors, and seniors within the multimodal communications. The students' storytelling aligns their contextual interactions, from which four

independent variables are involved to answer the hypotheses. These four independent variables—bodily gesture, articulation, facial expression, and voice loudness profoundly designate a positive and significant attribution toward students' storytelling performance. The multiple determinant coefficients produce .349 of its equivalence for 34.9% impact toward students' contextual interactions. Paralinguistic features undertake the fundamental capacity in the case of freshmen, sophomores, juniors, and seniors through multimodal communication, such as bodily gesture, articulation, facial expression, and voice loudness.

This conclusion withdraws that bodily gesture and facial expression rank in moderate categories, while articulation and voice loudness show attributable categories. These multimodalities accomplish students' contextual interactions in storytelling which define physical, psychological, and social matters through the spoken and nonverbal modes assessment. a through storytelling performance. This study suggests further research on paralinguistic features to widen and implement nonverbal modes. Lecturers are conditionally available to adopt and adapt options provided by the ELAN software, such as annotation, media synchronization, transcription, segmentation, and interlinearization modes, to investigate and analyze relevant subjects. These paralinguistic features empirically address the multimodal communication modes to improve teaching and learning activities.

Acknowledgement

This study was funded by DRTPM Kemdikbudristek, Republic of Indonesia under the research scheme for Penelitian *Desentralisasi Unggulan Perguruan Tinggi* (PDUPT), under contract number: PDUPT.DIKTI.02/UST/LP2M/K/VI/2022.

References

Aihara, R., Takiguchi, T., & Yasuo Ariki, Y. (2015). Individuality-preserving voice conversion for articulation disorders using phoneme-categorized. *ACM Transactions on Accessible Computing*, *6*(4), 1-17. https://doi.org/10.1145/2738048

Ali, A. Z. M., Segaran, K., & Hoe, T. W. (2015). Effects of verbal components in 3D talking-head on pronunciation learning among non-native speakers. *Educational Technology & Society*, *18*(2), 313-322.

Asadi, S., Wexler, A. S., Cappa, C. D., Barreda, S., Bouvier, N. M., & Ristenpart, W. D. (2019). Aerosol emission and super emission during human speech increase with voice loudness. *Scientific Reports*, *9*(2348), 1-11. https://doi.org/10.1038/s41598-01938808z

Behera, A., Matthew, P., Keidel, A., Vangorp, P., Fang, H., & Canning, S. (2020). Associating facial expressions and upper-body gestures with learning tasks for enhancing intelligent tutoring systems. *International Journal of Artificial Intelligence in Education*, *30*, 236-270. https://doi.org/10.1007/s40593-020-00195-2

Bursic, S., Boccignone, G., Ferrara, A., Alessandro D'Amelio, A., & Lanzarotti, R. (2020). Improving the accuracy of automatic facial expression recognition in speaking subjects with deep learning. *Applied Sciences*, *10*, 1-15. https://doi.org/10.3390/app10114002

Byun, S., Yoon, S., & Jung, K. (2020). Comparative studies on machine learning for paralinguistic signal compression and classification. *The Journal of Supercomputing*, *76*(1), 8357-8371. https://doi.org/10.1007/s11227-020-03346-3

Cabrera, L., Lorenzi, C., & Bertoncini, J. (2015). Infants discriminate voicing and place of articulation with reduced spectral and temporal modulation cues. *Journal of Speech, Language, and Hearing Research*, *58*(3), 1033-1042. https://doi.org/10.1044/2015_JSLHR-H-14-0121

Cánovas, C. P., Javier Valenzuela, J., Carrión, D. A., Olza, I., & Ramscar, M. (2020). Quantifying the speech-gesture relation with massive multimodal datasets: Informativity in time expressions. *PLoS ONE*, *15*(6), 1-18. https://doi.org/10.1371/journal.pone.0233892

Caridakis, G., Wagner, J., Raouzaiou, A., Lingenfelser, F., Karpouzis, K., & Andre, E. (2013). A cross-cultural, multimodal, affective corpus for gesture expressivity analysis. *Journal of Multimodal User Interfaces*, *7*, 121-134. https://doi.org/10.1007/s12193012-0112-x

Chanes, L., Wormwood, J. B., Betz, N., & Barrett, L. F. (2018). Facial expression predictions as drivers of social perception. *Journal of Personality and Social Psychology*, *114*(3), 380-396. https://doi.org/10.1037/pspa0000108

ELAN. (2021, July 6). *ELAN–Previous releases*. Nijmegen: Max Planck Institute for Psycholinguistics. https://archive.mpi.nl/tla/elan/previous

Esling, J. H., Moisik, S. R., Benner, A., & Crevier-Buchman, L. (2019). *1 – Voice and voice quality*. Cambridge Core: University of Sussex Library https://doi.org/10.1017/9781108696555.001

Gosztolya, G., & Tóth, L. (2018). A feature selection-based speaker clustering method for paralinguistic tasks. *Pattern Analysis & Applications*, *21*(1), 193-204. https://doi.org/10.1007/s10044-017-0612-0

Griethuijsen, R. A. L. F., Eijck, M. W., Haste, H., Brok, P. J., Skinner, N. C., Mansour, N., Gencer, A. S., & BouJaoude, S. (2014). Global patterns in students' views of science and interest in science. *Research in Science Education*, *45*(4), 581-603. https://doi.org/10.1007/s11165-014-9438-6

Gupta, R., Audhkhasi, K., Lee, S., & Narayanan, S. (2016). Detecting paralinguistic events in audio stream using context in features and probabilistic decisions. *Computer Speech and Language*, *36*, 72-92. https://doi.org/10.1016/j.csl.2015.08.003

Haj, M. E., Daoudi, M., Gallouj, K., Moustafa, A. A., & Nandrino, J-. L. (2018). When your face describes your memories: Facial expressions during retrieval of autobiographical memories. *Reviews in the Neurosciences*, *29*(8), 861-872. https://doi.org/10.1515/revneuro-2018-0001

Ishi, C. T., Mikata, R., & Ishiguro. (2020). Person-directed pointing gestures and interpersonal relationship: Expression of politeness to friendliness by android robots. *IEEE Robotics and Automation Letters*, *5*(4), 6081-6088. https://doi.org/10.1109/LRA.2020.3011354

Kano, T., Takamichi, S., Sakti, S., Neubig, G., Toda, T., & Nakamura, S. (2018). An end-to-end model for cross-lingual transformation of paralinguistic information. *Machine Translation*, *32*(2), 353-368. https://doi.org/10.1007/s10590-018-9217-7

Kaňok, M., & Novotný, M. (2019). Effect of age and gender on articulation of voiced and voiceless stop consonants in Czech. *Clinician and Technology*, *49*(3), 97-101. https://doi.org/10.14311/CTJ.2019.3.05

Kita, S., Alibali, M. W., & Chu, M. (2017). How do gestures influence thinking and speaking? The gesture-for-conceptualization hypothesis. *Psychological Review*, *124*(3), 245-266. http://doi.org/10.1037/rev0000059

Kita, S. (2009). Cross-cultural variation of speech-accompanying gesture: A review. *Language and Cognitive Processes*, *24*(2), 145-167. https://doi.org/10.1080/01690960802586188

Kozminska, K. (2018). Intonation, identity and contact-induced change among Polish-speaking migrants in the U.K. *Journal of Sociolinguistics*, *23*(1), 29-53. https://doi.org/10.1111/josl.12318

Kwon, A. Y., Vallotton, C. D., Kiegelmann, M., & Wilhelm, K. H. (2018). Cultural diversification of communicative gestures through early childhood: A comparison of children in English-, German-, and Chinese-speaking families. *Infant Behavior and Development*, *50*, 328-339. https://doi.org/10.1016/j.infbeh.2017.10.003

Labuschagne, P. A. (2017). Non-verbal and paralinguistic political communication in sport: An analysis of images of Springbok rugby. *Communication: South African Journal for Communication Theory and Research*, *43*(1), 80-102. https://doi.org/10.1080/02500167.2017.1307243

Leach, S., & Weick, M. (2020). When smiles (and frowns) speak words: Does power impact the correspondence between self-reported affect and facial expressions? *British Journal of Psychology*, *111*(1), 683-701. https://doi.org/10.1111/bjop.12433

Leung, Y., Oates, J., & Chan, S. P. (2018). Voice, articulation, and prosody contribute to listener perceptions of speaker gender: A systematic review and meta-analysis. *Journal of Speech, Language, and Hearing Research*, *61*(2), 266-297. https://doi.org/10.1044/2017_JSLHR-S-17-0067

Li, H., Yuan, Y. C., Bazarova, N. N., & Bell, B. S. (2018). Talk and let talk: The effects of language proficiency on speaking up and competence perceptions in multinational teams. *Group & Organization Management*, *44*(5), 953-989. https://doi.org/10.1177/1059601118756734

Li, K., Jin, Y., Akram, M. W., Han, R., & Chen, J. (2019). Facial expression recognition with convolutional neural networks via a new face cropping and rotation strategy. *The Visual Computer*, *36*, 391-404. https://doi.org/10.1007/s00371-019-01627-4

Lierde, K. M. V., Letter, M. D., Vermeersch, H., Roche, N., Stillaert, F., Lemmens, G., Peeters, P., Rogiers, X., Blondeel, P., & Corthals, P. (2015). Longitudinal progress of overall intelligibility, voice, resonance, articulation and oromyofunctional behavior during the first 21 months after Belgian facial transplantation. *Journal of Communication Disorders*, *54*, 42-56. https://doi.org/10.1016/j.jcomdis.2014.09.001

Lin, Y-. L. (2020). A helping hand for thinking and speaking: Effects of gesturing and task planning on second language narrative discourse. *System*, *91*, 1-10. https://doi.org/10.1016/j.system.2020.102243

Mortaji, L. (2018). Effects of sustained impromptu speaking and goal setting on public speaking competency development: A case study of EFL college students in Morocco. *English Language Teaching*, *11*(2), 82-98. http://doi.org/10.5539/elt.v11n2p82

Myers, B. R., & Finnegan, E. M. (2015). The effects of articulation on the perceived loudness of the projected voice. *Journal of Voice*, *29*(3), 3909-390.e15. https://doi.org/10.1016/j.jvoice.2014.07.022

Nandamudi, S., & Scherer, R. C. (2019). Airflow vibrato: Dependence on pitch and loudness. *Journal of Voice*, *33*(6), 815-830. https://doi.org/10.1016/j.jvoice.2018.05.007

Nilsen, E. S., Rints, A., Ethier, N., & Moroz, S. (2016). Mother–child communication: The influence of ADHD symptomatology and executive functioning on paralinguistic style. *Frontiers Psychology*, *7*(1203), 1-11. https://doi.org/10.3389/fpsyg.2016.01203

Nugrahaningsih, N., & Porta, M. (2019). Soft biometrics through hand gestures driven by visual stimuli. *ICT Express*, *5*(2), 94-99. https://doi.org/10.1016/j.icte.2018.04.014

Park, Y., & Stepp, C. E. (2018). The effects of stress type, vowel identity, baseline f0, and loudness on the relative fundamental frequency of individuals with healthy voices. *Journal of Voice*, *33*(5), 603-610. https://doi.org/10.1016/j.jvoice.2018.04.004

Pérez-Espinosa, H., Martínez-Miranda, J., Espinosa-Curiel, I., Rodríguez-Jacobo, J., & Avila-George, H. (2017). Using acoustic paralinguistic information to assess the interaction quality in speech-based systems for elderly users. *International Journal of Human-Computer Studies*, *98*, 1-13. https://doi.org/10.1016/j.ijhcs.2016.09.013

Pichel, B. (2016). From facial expressions to bodily gestures: Passions, photography and movement in French 19th-century sciences. *History of the Human Sciences*, *29*(1), 27-48. https://doi.org/10.1177/0952695115618592

Šebesta, P., Trebický, V., Fialová, J., & Havlícek, J. (2019). Roar of a champion: Loudness and voice pitch predict perceived fighting ability but not success in MMA fighters. *Frontiers Psychology*, *10*(859), 1-9. https://doi.org/10.3389/fpsyg.2019.00859

Sinha, S., Jain, A., & Agrawal, S. S. (2017). Empirical analysis of linguistic and paralinguistic information for automatic dialect classification. *Artificial Intelligence Review*, *51*(2), 647-672. https://doi.org/10.1007/s10462-017-9573-3

Slepian, M. L., & Carr, E. W. (2019). Facial expressions of authenticity: Emotion variability increases judgments of trustworthiness and leadership. *Cognition*, *183*, 82-98. https://doi.org/10.1016/j.cognition.2018.10.009

So, W-. C., & Wong, K. -Y. M. (2017). Tracing the development of spatially modulated gestures in the manual modality in nonsigning Chinese-speaking children. *Applied Psycholinguistics*, *39*(3), 527-544. https://doi.org/10.1017/S0142716417000510

Sumekto, D. R., & Setyawati, H. (2020). Revealing lecturer's paralinguistic attribution: How the visual manner contributes to students' non-cognitive skills. *Indonesian Journal of Applied Linguistics*, *9*(3), 559-571. https://doi.org/10.17509/ijal.v9i3.23206

Sumekto, D. R., Taufiqulloh, T., Setyawati, H., Hikmah, S., & Ghozali, I. (2021). Understanding the paralinguistic features disclosure depicted in the lecturer's visual modes of writing class instruction. *3L: Language, Linguistics, Literature*, *27*(4), 173-192. https://doi.org/10.17576/3L-2021-2704-13

Thacker, J. M., Chambers, C. G., & Graham, S. A. (2018). Five-year-olds' and adults' use of paralinguistic cues to overcome referential uncertainty. *Frontiers Psychology*, *9*(143), 1-10. https://doi.org/10.3389/fpsyg.2018.00143

Tooley, K. M., Konopka, A. E., & Watson, D. G. (2018). Assessing priming for prosodic representations: Speaking rate, intonational phrase boundaries, and pitch accenting. *Memory & Cognition*. *46*(4), 625-641. https://doi.org/10.3758/s13421-018-0789-5

Trujillo, J. P., Vaitonyte, J., Simanova, I., & Özyürek, A. (2019). Toward the markerless and automatic analysis of kinematic features: A toolkit for gesture and movement research. *Behavior Research Methods*, *51*, 769-777. https://doi.org/10.3758/s13428-01810868

Uştuk, Ö., & Aydın, S. (2016). The effects of the use of paralinguistic cues on foreign language anxiety among English as a foreign language speaker. *Innovation in Language Learning and Teaching*, *12*(3), 289-302. https://doi.org/10.1080/17501229.2016.1211133

Weisbuch, M., Grunberg, R. L., Slepian, M. L., & Ambady, N. (2016). Perceptions of variability in facial emotion influence beliefs about the stability of psychological characteristics. *Emotion*, *16*(7), 957-964. https://doi.org/10.1037/emo0000123

Wessel-Tolvig, B., & Paggio, P. (2016). Revisiting the thinking-for-speaking hypothesis: Speech and gesture representation of motion in Danish and Italian. *Journal of Pragmatics*, *99*, 39-61. https://doi.org/10.1016/j.pragma.2016.05.004

Wood, A., Martin, J. D., Alibali, M. W., & Niedenthal, P. M. (2018). A sad thumbs up: Incongruent gestures and disrupted sensorimotor activity both slow processing of facial expressions. *Cognition and Emotion*, *33*(6), 1-14. https://doi.org/10.1080/02699931.2018.1545634

Wu, H., Zhang, S., Liu, J., Qiu, J., & Zhang, X. (2018). The gesture disagreement problem in free-hand gesture interaction. *International Journal of Human–Computer Interaction*, *35*(3), 1-13. https://doi.org/10.1080/10447318.2018.1510607

Zarate, J. M., Tian, M., Woods, K. J. P., & Poeppel, D. (2015). Multiple levels of linguistic and paralinguistic features contribute to voice recognition. *Scientific Reports*, *5*(1), 1-9 https://doi.org/10.1038/srep11475