

Discussing in Zoom:
Investigating EFL students' self-monitoring and reactions to feedback

Shoko OTAKE, Miwa MORISHITA, Mayuko MATSUOKA

Abstract

Student group discussions are an important way of fostering communication in an EFL classroom. However, EFL Japanese students have difficulty with skills regarding Interactional Competence (IC), especially those that involve turn-taking. Thus, there is a need for classrooms to give quantitative feedback about IC skills to help with students' discussions. In previous studies we used an online video conferencing tool that shows students' speech activity in real-time. We found students can monitor their performance with the help of the website. In this current study, our aim was to see if students can still monitor their participation on Zoom, which does not show performance in real-time. Here, Japanese EFL college students participated in online group discussion sessions. The discussions were analyzed using a program developed by Hylable Inc. This program provided a feedback report for students that showed their individual performance. Overall, the quantitative data supported the students' own perception of their participation, though they were unable to reliably monitor others' reactions to their speech. Furthermore, students showed a positive reaction towards the feedback reports, and suggested their own plans for improvement. Future research may focus on individual differences in English skill and personality in order to foster motivation in a group discussion.

Key words: EFL, student discussion, Zoom, feedback, self-monitoring

1. Introduction

In a Japanese EFL conversation classroom, students need to practice various communication strategies. In some cases, this takes the form of oral presentations, where students need to understand logical presentation structure, and practice how to deliver their speech to their audience in a meaningful way. In other cases, however, students need to learn how to communicate in a less structured environment, such as a group discussion. However, even though teachers attempt to encourage communication and discussion in the classroom, previous research (Yanagi & Baker, 2015) has suggested that Japanese EFL students have the most difficulty in participating in class discussions when they study abroad.

1.1 Student discussion and interactional competence (IC) skills

Why do students have so much difficulty with discussions? Hauser (2009) suggests that this may come from the fact that Japanese students are used to a “monologic” style of discussion. Say that student A, student B, and student C have a discussion. A “monologic” style suggests, opinions are given in order, going from A to B to C. There would be no interruptions, no ideas are challenged in the strong sense, and students would not jump in to help another student develop their ideas. B and C may be reluctant to give ideas unless they are nominated by A to do so, like “So, A, what is your opinion?” All 3 students would give minimum responses (or backchanneling) to each other, such as nodding their head or saying “uh-huh”, just to be polite. Students used to this style are hit with a surprise when they study abroad. For example, student A can be interacting with student B, but student C may interrupt with a counterargument. In another part of the discussion, student B may negotiate and fix a misunderstanding that C has. Because of these differences, Yanagi & Baker (2015) found that students studying abroad in Australia found it very hard to break into a discussion and to turn-take.

Due to these difficulties, EFL Japanese students need to practice Interactional Competence (IC) skills. According to Young (2011) and Stone and Kershaw (2021), this concept includes many communicative behaviors such as agreeing/disagreeing, opening/closing statements, taking turns, and repairing problems. This type of competence should not rely on a single person, but should be shaped by all participants in a particular discussion. Research done by Pekarek-Doehler and Pochon-Berger (2015) have shown that these IC skills are not transferred from L1 to L2, but are something that students learn separately.

Over the years, researchers have studied how students may come to learn these skills. Stone and Kershaw (2021) found that students can be encouraged to take a look at recordings of their own discussions, and make a rubric identifying important IC skills. Overall, over a period of two months, students learned to open up the discussion topic in more collaborative ways. For example, in one statement, “Ok let’s start with best Japanese food, right?” a student learned to end a statement with ‘right?’ to give other members a chance to show understanding of the current topic. Other students would reply with “Yeah, okay.” to acknowledge that they understand where the discussion is heading. Other studies (Lam & Wong, 2000; Lam, 2009) have monitored discussion and provided students with various metacognitive strategies, leading to great improvements in discussion tasks.

1.2 Using technology to measure and assess IC skills

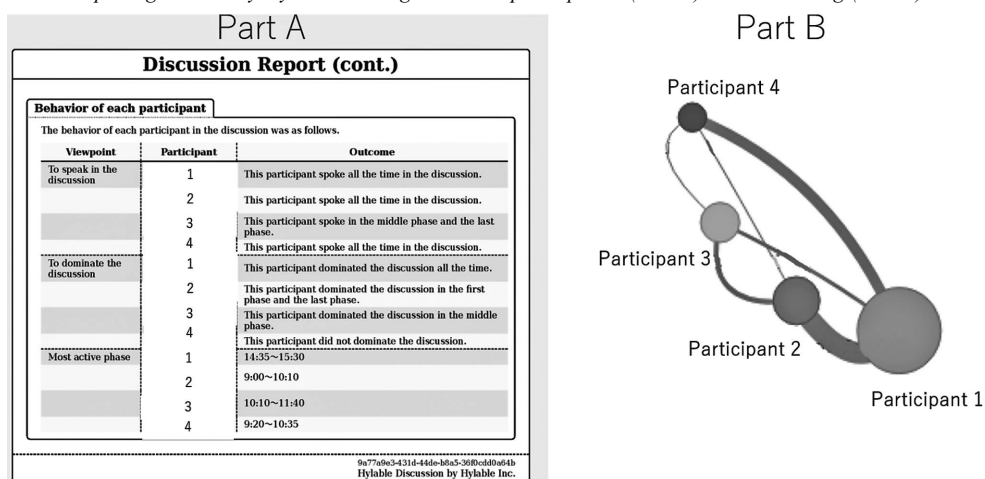
One question with IC skills, however, is how we can measure and monitor these skills in a reliable way. While using a rubric or using student self-reports are excellent solutions, some

language teachers may have different ideas about what IC skills are important. There may be individual differences in how students assess themselves, and some students may assess more honestly than others. Galaczi and Taylor (2018) suggest that using technology, such as online video conferencing tools, to monitor student speaking could be helpful, if used in conjunction with human assessment. Researchers like Warschauer (1995) have found that online settings can encourage students to participate equally in comparison to a face-to-face environment. While the reason for this is unclear, they suggest that perhaps online meetings give participants a sense of anonymity, and reduce social cues that make face-to-face conversations intimidating.

In our current study, we used Zoom (<https://zoom.us>) as a medium for students to do and record online group discussions. This was because the class was being held online on Zoom on a weekly basis, and students could easily be put into randomized 3-4 person groups using the Zoom function. We also used Hylable, a system that is developed by Hylable, Inc. to monitor discussion activity online. We chose to use this system because the system can quantitatively measure aspects of a group discussion. For example, the system can record the overall amount of speech produced by each student, how many times they turn-take, and how much students overlap with each other or keep silent. Importantly, this system provides each student with a report that visualizes their overall participation and behavior.

Figure 1 is an example of a report that students could get. This particular example shows the report of a four-person group. Parts of the report are edited in order to cut out personally identifiable information, such as student names and ID numbers. Part (A) of the report shows what part of the discussion students spoke the most in, who dominated the discussion, and times in the

Figure 1
Student reports generated by Hylable showing individual participation (Part A) and turn-taking (Part B)



discussion each student was most active in. Here, Participant 1 and 2 are the two students who seemed to dominate most parts of the discussion. Part (B) of the report depicts the turn-take pattern of the entire group. Each circle represents Participants 1 through 4, and the size represents how much individual students spoke. The thickness of lines connecting the circles represent how many turn-takes happened between students. For example, Participant 1 and 2 are connected with a very thick line, suggesting that many turn-takes happened between them. A turn-take here is defined as when one student speaks after another student. In our study, we asked students what they thought about this visual report, and how it may affect their interactions with their peers.

1.3 Research aims for our current study

Our first research aim was to investigate, “Can students reliably monitor their own performance, and others’ responses if Hylable feedback is not in real-time?” In a previous study (Otake, Matsuoka, Mizumoto & Morishita, 2021), we have used this Hylable system by having students log into a special online website. In this method, students saw moving graphs and charts indicating how much each person spoke over time. In other words, the feedback they received was in real-time. Overall, we found that this system helps students understand how much they are actually participating, as well as how much their classmates are responding to them. In their final ratings, we found that their overall contribution is based on both their overall amount of speech, as well as how much people respond to them. The results suggest that students were metacognitively aware of their speech and the relative responses of their surroundings. However, a limitation of this study was that we are not sure if students can actively monitor their behavior without the real-time feedback, which would be more representative of a standard online conversation. Therefore, in this current study students did their discussion in front of the class (via Zoom), and received a final report at the end. Then, we wanted to measure whether their own assessment of their performance was similar to the report produced by Hylable.

Our second research aim was to investigate, “What are student reactions towards the Hylable reports? Do they affect student performance?” In Otake et al. (2021), one of the issues we were unable to address was how students may (or may not) change their behavior according to the reports they received. Previous studies (Lam & Wong, 2000; Lam, 2009; Stone & Kershaw, 2021) have shown that reflecting over discussions can help students identify important IC skills. In this study, we asked students of the same class to do a group discussion over 2 sessions that were a week apart. In each session they were put into random groups of 3 or 4. For each session, students were given a series of discussion questions to think about in advance. They were explicitly told not to write out a script, but encouraged to do some background research and think about their own experiences.

Our third and final research aim was to get “insights from looking at student interactions.” As researchers (Galaczi & Taylor, 2018) suggest, it is important for human assessment to be used in conjunction with technological ones in order to truly look at IC skills. In our study, we created transcripts of the discussion, using the Zoom recordings. Since the participant group was too small to conduct stringent statistical analysis, we chose to pick up two groups of interest; one with the lowest amount of turn-taking, and the other with the highest amount of turn-taking. Then, we used LIWC-22 (Boyd, Ashokkumar, Seraj & Pennebaker et al., 2022), which is the abbreviation for “Linguistic Inquiry and Word Count.” LIWC is a type of text analysis software that counts and links words in a text to psychologically meaningful categories. For example, it can give general descriptives like words per sentence and types of function words. It can also assess emotionality of the writer (i.e., positive vs. negative) or even individual writing styles. Research (Pennebaker, Chung, Frazee, Laverne & Beaver, 2014) has suggested that individual differences in LIWC categories in college application essays are related to later college success. Here, we aimed to use this with the transcripts of two different groups, in order to capture any qualitative differences in their speech.

2. Method

2.1 Participants

Participants were 17 second-year university students who were studying English as a Foreign Language in the Kansai area. In terms of their English-speaking ability, they had a mean Versant score of 37.9 (range = 30-48), which is approximately CEFR A2 level and is a typical score of a Japanese university student. In their online English conversation classes, they went through a series of 6 units about social issues and did various oral activities based on the content. Throughout the semester, they practiced IC skill activities with each other, such as confirming ideas, agreeing/disagreeing, and using phrases to open/close a discussion topic. Students participated in the discussions and surveys as part of an end-of-semester assignment. They were asked to wear their own earphones or headset, in order to minimize background distractions.

2.2 Materials and Procedure

We used Zoom to have students participate in discussion, which is the same way students had been doing their regular class discussions and presentations. Students participated from home on their own computers. Overall, there were two recorded zoom sessions, each of which was separated by a week. Questionnaires were given to students using Google Forms, and most students completed them on their smartphones.

Prior to the discussion sessions, students were given three questions about “Gender”, as shown in

Table 1
Discussion Questions for Week 1 and 2

| Week 1: Jobs and Gender | Week 2: Women-only cars in Japan |
|---|---|
| (1) What kind of jobs are open to all genders? (2) What jobs should have more men? More women? (3) As a group, do you think jobs should be open to all genders? Or does it depend on the job? | (1) What are women-only cars? What are the rules? (2) Talk about your own experience with them. (3) Do you think women-only cars will disappear in the future? Why? |

Table 1. The questions were based on reading/listening assignments which they had already done. In the week prior to the discussion sessions, students were asked to take a look at the questions and think about what they would like to talk about. They were encouraged to do a recap of what they had already read about the topic if it would be easier, but were also encouraged to talk about their own experiences.

On Week 1, students came into their regular classroom zoom session and gave consent to participate in the study and to have their conversations recorded. The classroom teacher used the breakout room function on Zoom to randomly assign students to 3-4 person groups. First, groups were put into breakout rooms for 20 minutes to brainstorm and do additional research. After this, all groups came back into the main room. Each group presented their group discussion to the rest of the class for about 5-7 minutes, and the teacher commented on each groups' performance. Then, each group completed a questionnaire online. The questionnaire consisted of questions asking about (1) students' overall rating regarding their contribution (a scale of 1-5, where a 5 indicates lots of participation), (2) how much they thought others responded to them (a scale of 1-5, where a 5 indicates lots of responses), and (3) any comments about what they could have done better (given in Japanese).

On the following week (Week 2), students came into their regular zoom session again, and individually viewed their Hylable reports. Using the chat function on Zoom, they discussed what they should personally improve on for this second session. Similar to what was done on Week 1, the teacher again randomly assigned students to groups using the breakout room function. Students then took 20 minutes to brainstorm ideas for the Week 2 questions (see Table 1). Lastly, each group presented their group discussion to the rest of the class, and completed a final questionnaire. This final questionnaire was similar to the one given in Week 1.

Finally, we focused on two particular groups in Week 1 that showed contrasting performance (Group 2 and Group 3 from Week 1 had the lowest and highest numbers of turn-taking out of both weeks, respectively). In line with Boyd et al. (2022)'s recommendations, we removed everyday

filler words so they did not count as psychological state words (i.e., “you know” => “youknow”). We also removed slang words (“wanna” => “want to”), hyphens and ellipses.

3. Results

Table 2 shows the average contribution rating that students gave (1-5), collapsed across the first and second weeks. We analyzed the data using a generalized linear model with the predictors as (1) average proportion of time spent speaking, (2) turn taking (defined as when one student talks right after another student). We did not include the session (Week 1 or Week 2) as a predictor because it did not improve the model significantly ($p > .05$). We also did a separate analysis with students' ratings of how much turn-taking was going on (scale of 1-5), with the actual amount of turn-taking as the predictor variable. Table 3 shows the average turn-taking rating that students gave, collapsed across the first and second weeks. These analyses were conducted using the lme4 package in R (Bates et al, 2009).

For the first analysis, we found a significant main effect of proportion of time spent speaking, such that participants tended to give higher contribution ratings if they spent more of the time speaking ($b = 2.31$, $t(31) = 2.18$, $p < .05$). This was true regardless of whatever session students were taking part in. However, turn-taking did not significantly predict contribution ratings ($p < .05$). In our second analysis, the amount of turn-taking did not predict students' perceptions of how much turn-taking was occurring ($p < .05$).

Table 2
Average Proportion of Time Spent Speaking and Average Times Turn-Taking (Standard Deviations in Parentheses) by Contribution Rating

| Contribution Rating | Time spent speaking | Turn-taking |
|---------------------|---------------------|---------------|
| 2 | .11 (.03) | 28.50 (17.68) |
| 3 | .25 (.13) | 65.87 (66.52) |
| 4 | .30 (.11) | 86.00 (83.89) |
| 5 | .33 (.12) | 48.20 (26.52) |

Table 3
Average Times of Actual Turn-Takes (Standard Deviations in Parentheses) by each Turn-Taking Rating

| Turn-taking rating | Actual turn-taking |
|--------------------|--------------------|
| 2 | 28.50 (12.50) |
| 3 | 65.87 (64.41) |
| 4 | 86.00 (79.59) |
| 5 | 48.20 (23.72) |

Second, we looked at students' responses to Week 1 Hylable reports, though they were not statistically analyzed. When students looked at their reports from Week 1, only 2 students mentioned that the report showed lower contribution than they had expected, while others thought it was exactly what was expected (7 students), higher than expected (5 students), or much higher contribution than expected (3 students). To the free response question, "What would you like to improve on in this second week?", students suggested that they should be giving verbal responses (rather than nodding) to others' comments (5 students), being flexible in their speech (6 students), reducing unnecessary pauses (1 student) and improving their overall English ability (5 students).

Finally, we used LIWC to do a basic analysis of words used in two particular groups of interest, the groups with lowest (Group 2) and highest (Group 3) turn-taking. Table 4 and Table 5 show an excerpt of the transcript from Group 2 and Group 3, respectively. Table 6 shows Group 2 and 3's word counts, average words per sentence, rate of usage of function words, and overall emotional tone. Emotional tone in LIWC is a summary variable out of 100 points, where higher points (especially above 50) suggest a more positive emotional tone and lower points (especially below 50) suggest a more negative emotional tone.

Table 4
Excerpt From Group 2

| Speaker | Group 2 (low turn-taking) transcript |
|---------|---|
| A | Yeah these days I have seen many women work at station's office station's employee so it is important to uh equal the station's employees all genders. |
| | So what do you think [Student B]? |
| B | Yeah when I was looking for job at bank the rate of men and women did not change much so I thought it was very important for men and women to be equal. |
| A | Ok and do you think occupations should have been specific genders? |
| B | (Silence) |

Note: Ellipsis, repeated words, and slang words are left in this excerpt.

Table 5
Excerpt from Group 3

| Speaker | Group 2 (high turn-taking) transcript |
|---------|---|
| X | Because...when a man brought my dishes uh....his hand is very dirty. |
| Y | Ok....hahaha. |
| X | Um...women have a cleaner image. |
| Z | Yeah...ok ok. |
| Y | It's really like...it made me like disappointed or something. |
| X | Yeah, unlucky for you. |
| Y | Yeah, actually I am a waiter at a restaurant...so yeah...I...I'm shocked. Sorry. |

Note: Ellipsis, repeated words, and slang words are left in this excerpt.

Table 6
Group 2 and Group 3 Word Counts, Average Words Per Sentence, and LIWC Category Words

| | Group 2 | Group 3 |
|---------------------|---------|---------|
| Word count | 330.00 | 373.00 |
| Words per sentence | 12.22 | 6.11 |
| % Articles | 4.24 | 5.09 |
| % Prepositions | 9.70 | 9.38 |
| % Personal pronouns | 9.71 | 9.92 |
| % Auxiliary verbs | 12.12 | 14.75 |
| % Adverbs | 7.27 | 5.36 |
| % Conjunctions | 8.79 | 5.30 |
| Emotional Tone | 76.62 | 39.79 |

Note: Emotional tone is a summary variable based on a 100-point scale.

4. Discussion

4.1 Discussion of the 3 research aims

Our study was designed to explore three different research aims, all of which focus on EFL students' discussions and how Hylable reports may or may not help them in later sessions. Our first aim was to ask, "Do students monitor how much they are contributing to discussion, and how much people respond?" In previous studies, we had found that students can keep track of their own participation reliably, because their contribution ratings could be predicted by how much they actually spoke. In our current study, we found the same result, across both Weeks 1 and 2. Unlike previous studies, however, we did not find any reliable relationship between turn-taking and contribution ratings.

Furthermore, there was no reliable relationship between the students' perceptions of turn-taking

and the actual times it occurred. One reason for this could be that our current study did not show students how much they were turn-taking in real-time, and that made it difficult for them to keep track of others' responses alongside thinking about what to say themselves. The other possibility is that in this current case, the relationship is not a linear one. For example, looking at Table 2 we can see that turn-taking increases as the contribution rating becomes higher; however, it drops for students who gave a contribution rating of "5". This could indicate that there is a limit in the amount of turn-taking other members of a group engage in, if one person is speaking for a long time. This could also be the reason we did not see a relationship between student perceptions of turn-taking and actual turn-taking that occurred (Table 3).

The second aim we had was to investigate, "What are student reactions towards the Hylable reports? Do they affect student performance?" Most students thought the report was accurate, or a slight overrepresentation of their performance in Week 1. They were also able to identify some places for improvement, such as responding to others, which we had already identified as a part of discussion they were not monitoring very well. However, this did not improve student performance, at least in the short term. Between the two sessions, we saw no reliable changes in the amount of individual speech or turn-taking. This was especially true because including "Week (1 or 2)" as the predictor yielded no significant improvements to the model.

The final aim we had was to get "insights from looking at student interactions." Overall, the issue with low turn-taking groups seemed to be the fact that their contributions seemed to be memorized, and so students were unable to clarify what they meant. For example, in Table 4, Student A makes a great effort to facilitate the discussion. Noticeably, Student B only speaks after being nominated, and gives a fairly long response. However, when Student A asks for a clarification, Student B is unable to give it and there is a long pause. This is a typical monologic style of discussion, much like Hauser (2009) has discussed. On the other hand, if we take a look at Table 5 (Group 3), we can generally see that each student's speech is much shorter. Instead, lots of their speech is intended to indicate understanding of each other's speech. For example, Student X says "women have a cleaner image", and Student Y jokingly shows his shock towards such a bias towards the gender of restaurant workers.

Some aspects from the LIWC also get at this difference as well (Table 6); Group 2 uses more conjunctions than Group 3, suggesting that their (perhaps memorized) sentences are much longer and complex. On the other hand, Group 3 seems to use more auxiliary verbs, which have been associated with more narrative language styles (Jurafsky, Ranganath & McFarland, 2009). This may be in part because some Group 3 students shared their own stories related to the topic. Group 3 students also seemed to use more negative emotion words than Group 2, because they were discussing negative life events related to gender. In the excerpt, they use negative words like

“dirty”, “shock” and “unfortunate”.

4.2 Limitations and future directions

There are, of course, many limitations in this study that need to be addressed to see how to measure and improve IC skills in Japanese ESL students. In this study we failed to see any behavioral differences between Week 1 and 2. Part of the reason can be seen in student interactions (for the low turn-taking group), where some students show valiant efforts to keep the conversation going, whereas others do not see the point of interacting more than necessary. As Young (2011) pointed out, the success of interactions depends on all members of a group. One way to increase motivation is perhaps to have students think of the discussion questions by themselves, as Stone and Kershaw (2021) have done. In order for this to be a meaningful exercise, the teacher or the students' peers could help them generate questions that are open-ended. Another way is to gamify some student interactions, so that developing others' ideas or demonstrating disagreement gets them a certain number of points.

Another limitation in this study is the fact that it involves a small group of students, which made it difficult to make reliable comparisons between groups and individuals. By assessing a larger group of students from different classes, future studies could look at more reliable group dynamics and individual differences. For example, it could be beneficial for students if we group students with similar English-speaking ability (instead of randomly, as in this study). In our current study, we have found that when students contribute a lot to the discussion, the amount of turn-taking can reach a plateau. This may mean that there needs to be a balance between how much one speaker (who presumably could have high oral English skills) talks, and other students respond. Balancing the English oral skill levels in each group could leave room for students to figure out what a good balance is between their own speaking and allowing other students to speak.

Furthermore, looking at students' individual personalities could help students engage in their discussions. For example, in Zhang, Chen, and Liu (2020)'s study, they found that students' Big-Five personality traits have a strong effect on their affective engagement in online discussions. For example, they found that a group with members high in neuroticism should receive some extra attention from the teacher to avoid extra conflicts. In other cases, levels of conscientiousness and agreeableness had to vary greatly between group members in order for the discussion to move along. In a meta-analysis, Koutsoumpis et al. (2022) have found that LIWC categories and Big-Five personality traits are related as well. By assessing a larger student group with Hylable, LIWC, and the Big-Five, we may get better insights on how teachers could group students with certain Big-Five traits in a way that benefits discussion.

Overall, the results of the study provide an example for how Japanese EFL students interact with

each other in a classroom discussion. Even after a semester of discussion practice, IC skill instruction, and feedback in the current study, many students find difficulty moving from a “monologic” style of discussion to a less structured style of discussion. Since having a successful discussion is a group effort, future studies may benefit by finding a good balance of individual English ability and personalities in discussion members.

5. Additional notes

Parts of this manuscript were based on our presentation given at the 42nd Thailand TESOL conference (January 27-28, 2023).

6. Acknowledgements

This study was supported by a grant from Kobe Gakuin University for special overseas research. We thank Hylable, Inc. and Dr. Takeshi Mizumoto for their technical support and feedback on earlier versions of this manuscript.

References

- Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., Dai, B., Scheipl, F. & Green, P. (2009). Package ‘lme4’. URL <http://lme4.r-forge.r-project.org>.
- Boyd, R. L., Ashokkumar, A., Seraj, S., & Pennebaker, J. W. (2022). The development and psychometric properties of LIWC-22. *Austin, TX: University of Texas at Austin*, 1-47.
- Doehler, S. P., & Pochon-Berger, E. (2015). The development of L2 interactional competence: Evidence from turn-taking organization, sequence organization, repair organization and preference organization. *Usage-based perspectives on second language learning*, 30, 233.
- Galaczi, E., & Taylor, L. (2018). Interactional competence: Conceptualisations, operationalisations, and outstanding questions. *Language Assessment Quarterly*, 15(3), 219-236.
- Green, C. F., Christopher, E. R., & Lam, J (1997). *Developing discussion skills in the ESL classroom*, *ELT Journal*, 51(2), 135-143.
- Jurafsky, D., Ranganath, R., & McFarland, D. (2009, June). Extracting social meaning: Identifying interactional style in spoken conversation. In *Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics* (pp.638-646).
- Koutsoumpis, A., Oostrom, J. K., Holtrop, D., Van Breda, W., Ghassemi, S., & de Vries, R. E. (2022). The kernel of truth in text-based personality assessment: A meta-analysis of the relations between the Big Five and the Linguistic Inquiry and Word Count (LIWC). *Psychological Bulletin*, 148(11-12), 843.

- Lam, W. Y. K. (2009). Examining the effects of metacognitive strategy instruction on ESL group discussions: A synthesis of approaches. *Language Teaching Research*, 13(2), 129-150.
- Lam, W. Y. K., & Wong, J. (2000). The effects of strategy training on developing discussion skills in an ESL classroom, *ELT Journal*, 54(3), 245-255.
- Otake, S., Matsuoka, M., Morishita, M., & Mizumoto, T. (2021), What does a good discussion consist of? Insights from quantitative analysis of Japanese EFL students' online discussions, IALLT Conference, Virtual.
- Pennebaker, J. W., Chung, C. K., Frazee, J., Lavergne, G. M., & Beaver, D. I. (2014). When small words foretell academic success: The case of college admissions essays. *PloS one*, 9(12), e 115844.
- Stone, P. & Kershaw, M. (2021). *Developing practices for opening a classroom discussion* [Doctoral dissertation, Gakushuin University].
- Warschauer, M. (1995). Comparing face-to-face and electronic discussion in the second language classroom. *CALICO Journal*, 7-26.
- Yanagi, M. & Baker, A. A. (2015). Challenges experienced by Japanese students with oral communication skills in Australian universities. *TESOL Journal*, 7(3), 621-644.
- Young, R. F. (2011). Interactional competence in language learning, teaching, and testing. *Handbook of research in second language teaching and learning*, 2, 426-443.
- Zhang, X., Chen, G., & Xu, B. (2020). The influence of group Big-Five personality composition on student engagement in online discussion. *International Journal of Information and Education Technology*, 10(10), 744-750.