



An analysis of turn-taking behaviors of Japanese learners of English in videoconferencing discussions

 Jonathan Hennessy

Toyo University, Japan

hennessy.jonathan@gmail.com

To successfully interact in a new language, learners must effectively take turns and manage the floor. Expectations and rules for turn-taking can differ between languages, making this more than a question of grammatical proficiency and vocabulary acquisition. In addition, the increased use of videoconferencing software in education and beyond makes learning to manage the floor in virtual spaces a useful skill for everyone, including language learners. To analyze the turn-taking behaviors of first language (L1) speakers of Japanese using English in videoconferencing interactions, three classes of first-year university students with intermediate level English skills were selected to have their in-class videoconferencing discussions recorded. Three of the recorded discussions were selected and analyzed to identify turn-taking behaviors and to analyze the techniques that influenced the speakers' ability to succeed. Participants were observed to leave long gaps between speakers at turn transitions and rarely extended discussion topics to include multiple turns per speaker. There were three behaviors that were observed that helped some participants to reduce the gap between speakers. Clearly marking the end of a turn, energetic use of verbal backchannels, and comfort with unintentional overlap all seemed to be correlated with improved transition speed. Increased use of questions did lead to some topics being discussed beyond a single turn per speaker, but the difference was relatively small.

Keywords: videoconferencing, online interaction, turn-taking, discourse analysis

Introduction

Videoconferencing software has been in use for more than two decades (Reidl, 2021) and has been used for second language education (Goodfellow et al., 1996; Tecador Cabrero, 2013). As a result of the COVID-19 pandemic, use of

videoconferencing software has rapidly increased (Nikou, 2021; Reidl, 2021). This trend appears in education, with options such as Zoom, Microsoft Teams, and WebEX being used for online instruction (Nikou, 2021), and in other contexts (Riedl, 2021). Even as COVID restrictions are reduced and face-to-face interactions increase, videoconferencing may remain a significant tool for many people (Marshall et al., 2020; Nikou, 2021).

If videoconferencing continues to be an important method of interaction, learning to effectively interact with others through videoconferencing software is a valuable skill to acquire. Floor management, the method of organizing and proceeding through spoken interactions, and turn-taking dynamics in face-to-face conversations are challenging for second language learners (Cook, 1989; Wong & Waring, 2010). Additionally, these elements of interaction are managed differently in videoconferencing conversations (e.g. Boland et al., 2021; Reidl, 2021). Failures in turn-taking and floor management can lead to a complete inability to participate (Wong & Waring, 2010). This suggests that developing these skills to allow for effective videoconferencing interactions could be beneficial for language learners. This paper analyzes three videoconferencing discussions among first-year Japanese university students using English as a foreign language. The analysis describes the behaviors and techniques used to facilitate floor management and turn-taking during these online discussions and attempts to identify which methods were effective for the participants involved. The analysis makes use of the Conversation Analysis methodology and uses excerpts from the transcriptions of these discussions to attempt to understand the reasoning behind the interactive behaviors of the observed speakers (Wong & Waring, 2010).

Literature review

Turn-taking

Taking turns is one of the most fundamental and important features of spoken interaction (Tanaka, 1999; Wilson & Zimmerman, 1986; Wong & Waring, 2010). Turns in an interaction are clearly organized (Sacks et al., 1974), with the organization and use of turn-taking mechanisms being managed by participants in a given interaction (Ford & Thompson, 1996; Sacks et al., 1974). Turns are managed with high precision and are clearly connected (Cook, 1989; Sacks et al., 1974). Most turns have no overlap or gap between speakers (Sacks et al., 1974) and silence between turns is rare and brief (Wilson & Zimmerman, 1986). Overlap between speakers is infrequent (Cook, 1989; Dörnyei & Thurrell, 1994; Sacks et al., 1974) and often has semantic significance (Cook, 1989). Interactions follow rules to determine who has the right to speak and to help speakers avoid speaking simultaneously (Dörnyei & Thurrell, 1994; Sacks et al., 1974). When transitioning between speakers, techniques are generally used to determine who has the right to speak, with most transitions being some variation of self-selection, where a participant selects themselves as the next speaker, or a variation of the current speaker choosing who should speak next (Sacks et al., 1974).

Turns can be defined both in terms of the mechanics of the language (Ford & Thompson, 1996; Furo, 2001; Sacks et al., 1974) and in terms of the interactional value of turns as they are negotiated by a group (Barraja-Rohan, 2011; Celce-Murcia, 2007; Wong & Waring, 2010). Mechanical definitions consider turns as being made up of “turn-constructual units” (TCU), which are sections of an utterance that could be considered a complete speaking turn (Sacks et al., 1974). Sacks et al. (1974) also posited the idea of transition-relevance places (TRP) as places where speaker changes were possible. This has occasionally been adjusted to complex transition-relevance places (CTRP), where some combination of grammatical, semantic, and intonational end points occur simultaneously, signaling where turn transitions are not only possible, but likely (Clancy et al., 1996; Ford & Thompson, 1996; Furo, 2001). Rather than focusing on structural elements, interactional definitions focus on how each group of speakers negotiates the rules they use for turn management (Barraja-Rohan, 2011; Celce-Murcia 2007; Wong & Waring, 2010).

The rules governing taking turns and managing the floor are varied between languages and cultures (Cook, 1989; Du-Babcock, 1999; Dörnyei & Thurrell, 1994). When learning a new language, it can be challenging to learn appropriate techniques for taking turns (Cook, 1989; Ryan & Forrest, 2021). Challenges with turn-taking can lead to speakers struggling to participate when interacting with others who are more comfortable with the norms of that language (Cui, 2014; Lee, 2009; Morita, 2000; Ryan & Forrest, 2021) and speakers may find themselves unable to get a chance to speak at all (Wong & Waring, 2010). Longer gaps when a speaker is expected to speak can also lead to negative perceptions of fluency (van Os et al., 2020). To succeed in using a new language, learners must learn how to take, pass, and hold the floor according to the rules of the new language (Celce-Murcia, 2007). Variation from the norms of L1 users of a language does not inherently indicate a failure in acquisition (e.g. Kachru, 1990) and it is not always necessary for instructors to force their students to adapt to a particular style of use (Harumi, 2001). However, the research cited here does suggest that some proficiency with target language floor management techniques can be beneficial to language learners.

Interlanguage: Turn-taking in English and Japanese

The use of pragmatic skills in a new language has been observed to be affected by language transfer and existing cultural expectations (e.g. Du-Babcock, 1999; Kasper, 1992; Saito & Beecken, 1997), including in Japanese learners of English (Du-Babcock & Tanaka, 2013; Takahashi & Beebe, 1987; Tanaka, 2006). English is frequently used for communication between L2 speakers of different linguistic backgrounds (e.g. Jenkins, 2009; Sowden, 2012) and variations in usage are not inherently problematic. As such, it is necessary to differentiate between behaviors that cause communication issues and those that do not. The influence of a learner’s previously learned languages has been cited as influencing the use of turn-taking techniques, which can in some cases lead to difficulties in effective communication (Cui, 2014; Harumi, 2001; Morita, 2000; Tanaka,

2008; Williamson, 2019; Young, 2018). Harumi (2001) surveyed and observed undergraduate English majors at a Japanese university and suggested transfer from Japanese as a reason that Japanese learners of English used silence in conversation more frequently than L1 English speakers. Kitamura (2001) noted that in observations of Japanese speakers of English that they strongly avoided overlap in casual conversation and used silence as a method of repairing overlap. While this could suggest that we should expect more silence in conversations between L1 speakers of Japanese, Stivers et al. (2009) found speakers of Japanese to be comparatively fast at making turn transitions compared to a variety of other languages, including English. Tanaka (2006) observed and interviewed Japanese speakers of English at a Japanese subsidiary of an American company and detailed how they seemed to make use of L1 active listening techniques and cultural assumptions about when to speak out against an idea they disagreed with. Du-Babcock and Tanaka (2013) observed business meetings in English between people from Hong Kong and Japan and noted that the Japanese participants behaved in ways more commonly observed in Japanese interactions when expressing disagreements. Kern (2009) noted a varied level of comfort with interrupting other speakers among Japanese learners of English at a university in Japan. While these behaviors are not inherently problematic, the previous section of this paper notes that when turn-taking is affected, speakers may face interactional challenges. Tanaka (2008) observed this when he detailed a business interaction between L1 speakers of Japanese and French interacting in English. He argued that the use of L1 Japanese norms in the L2 interaction put the Japanese participants at a significant disadvantage in a negotiation.

There are different perspectives as to what first language and cultural factors influence a learner's interlanguage. Cui (2014) and Morita (2000) both found that language learners reported feeling uncomfortable using turn-taking strategies that they knew were appropriate in their new language, though neither study was focused specifically on L1 speakers of Japanese. Young (2018) suggests that the linguistic structure of Japanese impacts behavior around turn-transitions. Williamson (2019) suggests that cultural norms and pragmatic knowledge are the likely reasons for the differences in floor management.

English and Japanese speakers communicating in their first languages also approach turn-taking and floor management differently (Fujimoto, 2010; Furo, 2001; Kern, 2009; Tanaka, 1999). Differences have been observed in how listeners act while another person is holding the floor (Fujimoto, 2010; Furo, 2001; Kern, 2009; Tanaka, 1999), how speakers try to maintain their right to the floor (Furo, 2001), and how cultural rules impact turn order (Fujimoto, 2010; Furo, 2001).

In Japanese interactions, backchannels, when a listener responds to another speaker without taking the floor, are more common than in English interactions (Fujimoto, 2010; Furo, 2001; Maynard, 1986). Furo (2001) observed this when comparing discussions between friends for speakers of both Japanese and English during L1 interactions. Maynard (1986) studied American and Japanese university students and found that while the purpose of backchannels

seemed to be the same, they were more frequent with the Japanese students. L1 English speakers are also more likely to interrupt (Furo, 2001; Kern, 2009). Furo (2001) explained that in English interactions interruption can be perceived as a sign of participation whereas in Japanese avoiding interruption contributes to the current speaker's right to the floor. L1 English speakers frequently begin floor changes before a speaking turn has finished, which is less common in Japanese (Furo, 2001). Tanaka (1999) and Young (2018) both noted that the difference in grammar forces listeners to wait longer into an utterance to predict the intent of the turn, potentially contributing to this behavior.

When trying to maintain their right to the floor around CTRPs, Furo (2001) noted that Japanese speakers tend to elicit backchannels to maintain support, whereas English speakers may skip expected intonational end points or rush through an end point. Tanaka (1999) analyzed data from across a large section of L1 Japanese speakers including phone conversations, interactions at home, and business interactions. He noted that Japanese speakers also rush at the end of a TCU to maintain the floor and may additionally use particles that suggest a turn isn't finished to hold onto the floor.

There are also differences in the cultural rules about turn order. While Sacks et al. (1974) assert that turn order is not decided before an interaction, research into Japanese speakers suggests that this might not always be true. Fujimoto (2010) claimed that Japanese speakers may talk about the order of topics and turns before beginning a conversation. Other researchers have suggested that age and gender may also influence turn order in Japanese conversations (Fujimoto, 2010; Furo, 2001).

The variations found between Japanese and English interactions may offer some explanatory power for the behaviors that can be observed when L1 speakers of Japanese interact in English. Some variation in floor management should be expected compared to what is observed in L1 users of English. Understanding the features of L1 Japanese interaction can help differentiate between interactional features that are influenced by L1 norms and those that are affected by communicating with videoconferencing software.

Turn-taking in videoconference conversations

Videoconferencing interactions tend to have more gaps and pauses, and overlap can be a more difficult problem to solve compared to in face-to-face interactions (Boland et al., 2021; Gettliffe & Erkat, 2016; Kawashima et al., 2008; Nakaya & Okamoto, 2022; Smidt et al., 2017; Taylor, 2011). Taylor (2011) observed university students at a university in the UK and noted that participants perceived pauses in videoconferencing to be longer than in face-to-face interactions and that they found it difficult to predict when another person would speak. Kawashima et al. (2008) noted that the delay between turns was longer in videoconferencing conversations. Boland et al. (2021) found that the duration of pauses in Zoom conversations among university students was more than double that in face-to-face conversations. Gettliffe and Erkat (2016) noted that online foreign language tutors at an American university reported

an increased frequency in overlaps caused by difficulty in turn management. Smidt et al. (2017) found that unintentional overlaps in an online graduate course often caused all participants to stop speaking. Nakaya and Okamoto (2022) compared a face-to-face party of Japanese university students to a similar event held online and found that in the videoconferencing interaction, overlaps that caused a speaker to stop talking were more common than in face-to-face interactions. Tecador Cabrero (2013) suggested that the slow pace of online conversations can make it difficult to tell when a speaker wishes to hold the floor, leading to more pauses during turn transitions.

The issues found in floor management during videoconferencing interactions could partially be the result of the challenges inherent with communicating through body language while using videoconferencing software. Both English and Japanese language research have noted a lack of or increased difficulty of using gaze, eye contact, and body language (Goodfellow et al., 1996; Imai et al., 2022; Reidl, 2021; Taylor, 2011). Goodfellow et al. (1996) suggested that the inability to combine gaze with open questions can impact floor management in a conversation based on their observations of a videoconferencing interaction that was part of a professional English course. Imai et al. (2022) suggested that the ineffectiveness of eye contact and gaze leads to increased stoppages and gaps when changing speakers. Reidl (2021) similarly argued that videoconferencing conversations are more difficult to coordinate due to the reduced eye contact and difficulty of reading body language. Participants in videoconferencing conversations also seem to be aware of these issues, as Taylor (2011) reported that some speakers felt that the flow of dialogue was negatively affected by reduced eye contact.

Other features of videoconferencing interactions are less thoroughly discussed. Takao (1999) noted videoconferencing conversations between Japanese university students often included longer monologue style turns. Tanaka and Mori (2022) saw a shift to content at the cost of interpersonal connection with videoconferencing conversations between Japanese university students. The technology used in videoconferencing can also have an impact. Reidl (2021) stated that even a miniscule delay in audio can make videoconferencing feel more difficult than face-to-face interaction. Tanaka and Mori (2022) noted that despite full duplex audio making it possible to hear multiple speakers simultaneously, listeners found processing this simultaneous speech to be more difficult and adapted to avoid overlap. Gettliffe and Erkat (2016) noted that speakers would pause after a turn as a way to avoid technical problems impeding understanding.

With more experience, many speakers adapt to online conversations and turn-taking performance improves (Kawashima et al., 2008; Smidt et al. 2017; Taylor, 2011; Tecador Cabrero, 2013). Speakers also find different ways to help manage turns during videoconferencing interactions. The strategies varied between studies and included making use of hand movement around transitions (Nakaya & Okamoto, 2022) as well as leaving the microphone active to allow speakers to hear backchannels (Moorhouse et al., 2022). Thus, it is reasonable to expect groups to develop techniques for managing the floor when

using videoconferencing software, but different groups may choose different methods of doing so.

Research methods

Research design

The analysis in this paper is based on the Conversation Analysis framework as a way to analyze talk as it occurs naturally (Wong & Waring, 2010). The data presented were collected from a discussion test that was given as a regular part of an English discussion course. The analysis begins with “unmotivated looking” (Wong & Waring, 2010, p. 6) and as the data is observed, transcribed, and reviewed, areas of focus are identified and selected for analysis (ten Have, 2007; Wong & Waring, 2010).

Participants

The participants for this study were 11 first-year university students taking a required English discussion course at a private university in Tokyo. All participants were L1 Japanese speakers. The classes that were recorded were of the second highest of the four proficiency bands used by the university for sorting students. The university did not provide exact score levels for each class, but instructors are told that students in those classes have Test of English for International Communication (TOEIC) scores between 480 and 680. Generally, students assigned to this band at the university are capable of basic interaction on the topics they are asked to discuss without excessive interference from lack of proficiency. This is especially true towards the end of the semester after they have developed a rapport with their classmates and have learned techniques and phrases to help share their ideas. That said, it is likely that some variations between groups are at least partly a result of differences in linguistic proficiency.

Data collection

The recorded discussions all came from the final discussion test in the 13th lesson of a 14-lesson semester. This was chosen as test lessons were the most practical to record and featured the least interference from the instructor. Collecting data at the end of the semester gave the most opportunity for students to adapt to videoconferencing, which prior research has shown improves interaction (Kawashima et al., 2008; Smidt et al., 2017; Taylor, 2011; Tecador Cabrero, 2013). The discussions were held and recorded using the videoconferencing software Zoom. Zoom uses full duplex audio which allows for simultaneous speakers to remain audible, ensuring that backchannels and overlap would not be excessively problematic due to technical limitations.

Context of the recordings

The discussions were time limited based on the number of participants and each group was given identical topics. Participants were instructed that they could discuss whichever topics they wished with no requirement for topic order or completion. There was no intervention made, but students were scored on their use of skills taught in the class as well as their participation through holding the floor and asking questions. The assessment results are not a part of the analysis in this paper, but some explanation is provided as the use of skills targeted for assessment and previously taught skills did have some effect on the way that students managed their discussions.

The skills assessed in the recorded discussions included skills to confirm comprehension and repair any breakdowns as well as content focused skills. For the final test, which is analyzed in this paper, there were three content skills students were expected to make use of. The first was considering different viewpoints, where students were supposed to ask what other people might think of a topic and speculate on the thoughts and ideas of others. The second was asking for and providing advantages and disadvantages for their ideas to consider their own suggestions more deeply. The final skill was to ask for and provide sources for the information that they shared during the discussions. Students also made use of previously learned skills that were not targeted for assessment during this evaluation including asking for, giving, and supporting opinions as well as floor management skills including selecting topics for discussion, inviting others to speak, requesting the floor, and confirming whether the entire group was ready to move on from the current topic of discussion.

Data analysis

Eight total discussions from four classes were recorded and three were chosen for analysis. All discussions were reviewed to identify the target of analysis and to select specific discussions for detailed consideration. The three chosen recordings were selected as they had few to no technical issues, did not suffer from comprehension issues caused by background noise, and demonstrated a range of interactional behaviors.

The three discussions analyzed will be referred to as discussions A, B, and C. Discussion A featured three participants and had a 12-minute time limit whereas discussions B and C each had four participants and were 16 minutes. Discussion A included two female participants, referred to as Eri and Aki, and one male, referred to as Jun. In discussion B there were three female participants, referred to as Ami, Saki, and Yui, and one male participant, referred to as Ryo. Discussion C had two female participants, Mai and Yuna, and two males, Ken and Shun.

The discussions were analyzed using the video editing software, Lightworks, which allowed the 25 frame per second video to be analyzed with accuracy to 0.04 seconds. They were transcribed using the transcription key included in Appendix A, which is adapted with minor adjustments from Furo (2001). The analysis is done using a Conversation Analysis framework, meaning that the

analysis is focused primarily on how speakers interact, and the research questions are shaped as the interactions and transcriptions are examined (ten Have, 2007). Patterns of interactional behavior will be noted, and examples from the interaction will be presented and explained in order to analyze how these speakers managed their interactions.

While observing the discussions and reviewing the transcripts, participants' behavior around floor management and the pauses around speaker changes stood out as an area of interest. Specifically, the analysis will look at what Furo (2001) defines as floor taking speaker changes, when a speaker holds the floor to provide content to the discussion through a question or a statement. With this focus in mind, the transcripts were analyzed line-by-line to identify and explain the behaviors of the participants and interactional features that influenced floor management in each discussion. The research questions that will guide the analysis of the results are:

1. In what ways did the participants succeed and struggle at managing turns in their discussion?
2. What techniques did the participants use to manage the floor and facilitate turn-taking and how effective were these techniques?
3. Were there any other behaviors or occurrences that could contribute to the relative success or failure of turn-taking in the discussion?

Results and discussion

Basic floor management: Gaps and turn distribution

In all discussions, the participants were able to speak for the entire duration of the test with no communication issues large enough to significantly derail their interactions. No group stopped their discussion before time expired, Japanese was not required to assist in turn management, turn allocation was relatively balanced, and no participants failed to participate. However, in analyzing the videos, two unexpected behaviors were observed. The first was the frequent presence of long gaps between speakers at turn transitions. As was noted in the literature review, some gap is to be expected in videoconferencing conversations (Boland et al., 2021; Gettliffe & Erkat, 2016). However, these gaps were long enough to be considered a sign of difficulty and could create negative impressions of the participants' fluency (van Os et al., 2020). The second issue was a lack of cooperative interaction. Speakers generally spoke once per topic with minimal influence drawn from other speakers' turns and few elaborative questions.

Since the discussions were intended to be as organic as possible, many transitions were affected by factors other than the participants' ability to rapidly change speakers in a controlled environment. For example, some transitions were slowed down by participants laughing or giving extended positive non-linguistic responses to a speaker's comment. Marking this as a gap would not accurately represent the actual behavior of the participants. The same is true for transitions where gaps were clearly caused not by any issues with claiming

the floor, but because all participants were thinking about what they wanted to say and communicating that with thinking sounds. As a result, a full quantitative analysis of all transitions and gaps would not accurately represent the interactive behavior of the participants. To give a general impression of the speed of transitions, the transcriptions were analyzed and all cases where transitions occurred without any interference were noted, and the gap duration was recorded. Additionally, the low number of analyzed transitions means that a quantitative analysis would not be appropriate due to the sample size. The data are shared here to help justify and understand the excerpts used in the qualitative analysis that follows.

Discussion A had 19 transitions of this type with a mean duration of 4.22 seconds and a median of 3.16 seconds. Ten gaps were longer than three seconds while the rest were under two seconds. Of those shorter gaps, five occurred during a single incidence of negotiation for meaning between two students including three that were less than one second that could be considered to have no gap. Discussion B had 16 transitions to analyze with a mean gap duration of 5.15 seconds and a median of 4.5 seconds. Ten were longer than three seconds with the remainder lasting between one and two seconds. Discussion C had far more transitions without interference and 31 transitions were recorded with a mean duration of 1.85 seconds and a median of 1.24 seconds. Additionally, only six of the gaps were longer than three seconds, ten were less than one second, and 11 more were under two seconds. This is summarized in table 1 below.

Table 1. Duration of gaps without interference

Gap duration	<1 second	1-2 seconds	2-3 seconds	> 3 seconds	Total	Mean gap	Median gap
Discussion A	3	6	0	10	19	4.22	3.16
Discussion B	0	6	0	10	16	5.15	4.5
Discussion C	10	11	5	5	31	1.85	1.24

Qualitatively, the long gaps between speakers in discussions A and B suggested a hesitation to claim the floor. To illustrate this, two examples are provided below. These were selected as the gaps between turns were easily noticed without any clear reason for the silence. The following example from discussion A includes a gap in excess of five seconds with only backchannels to fill the space.

Jun: So, providing free health services for poor people is good

Jun: solution. /># {5.24}

Aki: *' Wow.

Eri: *' Yes.

Eri: " Um. I partly agree with you. />#

This next example, taken from discussion B, shows a speaker marking the end

of their turn with a question and being met with more than six seconds of silence before another participant took the floor.

Saki: We should put some restruction, restriction? /> [Uh restriction.
Yui: [*' Yes, yes.
Ryo: [*' Yes.
Saki: What do you think about my idea? /># {6.12}
Yui: " Can I say something? />

Both groups also had transitions with much shorter gaps when it was clear who was expected to speak next. This example from discussion A features a gap of only one second when the new speaker is answering a follow-up question.

Eri: I didn't know that. /># How do you know about that? /># {1.00}
Aki: I read a, I read Benesse website. />#

The next example from discussion B features a gap just longer than one second when the next speaker is selected by the current speaker.

Ryo: So if Japanese government raise taxes for rich people, they lost
Ryo: working motivation. /># So I think it is bad. />#
Ryo: How about you Saki? /># {1.08}
Saki: " Mmm, I think it is good. />#

Participants occasionally transitioned without a clear next speaker without excessively long gaps, but long transitions when the next speaker had to voluntarily claim the floor through self-selection were common in both discussion A and B. The extended gaps here are consistent with observations from previous research that observed increased gap duration between turns in videoconferencing interactions (e.g. Boland et al, 2021; Tecador Cabrero, 2013).

Discussion C also included some transitions with long gaps; however, short gaps were much more prevalent than in discussions A and B. In addition, these short gaps were not reliant on the next speaker being identified ahead of time. The example given here illustrates this point well, as it includes a simple transition with no interference and no techniques used to select the next speaker yet features a gap slightly below the average for simple transitions in this group.

Shun: So I think providing free childcare is very good for poor
Shun: parents. /># (01.76) Is there anything to add/>?# {01.08}
Mai: " Yes. Ahh, as Shun said, from the point of from single mothers' point
Mai: of view they have to work and raise children. /># It is so hard. />#

Given the delay in transmission introduced by videoconferencing (Gettliffe & Erkat, 2016; Reidl, 2021; Tanaka & Mori, 2022), gaps like this could be considered generally successful. Comparing the three discussions suggests that gaps in videoconferencing conversations can be reduced in duration with skill.

However, it is not clear from this data that all groups of speakers will achieve this.

The other noteworthy behavior was the lack of interaction between participants. Sacks et al. (1974) asserted that the relative distribution of turns is not decided before a conversation starts, however, in all three discussions participants generally held the floor exactly once per topic with little variation. Not counting negotiation of meaning, discussion A had only two deviations; once when one speaker asked a question to a previous speaker and one follow-up question asked to the current speaker yielding another turn for that current speaker. Discussion B also had one delayed question and one follow-up question. Additionally, in two of the four topics they discussed, some participants chose not to take the floor. It is reasonable for speakers to choose not to take the floor for every topic and over the course of the discussion it did not seem like participants were unable to speak when they wanted to. However, despite creating some variety in turn distribution, it does not seem appropriate to suggest that this in and of itself indicates freely varied turn distribution. Fujimoto (2010) claimed that turn-order in Japanese interactions may be more predictable than in interactions of L1 English speakers, so this may simply be a variation to be expected with L1 speakers of Japanese.

Discussion C was slightly more varied. In addition to negotiation for meaning there were three follow-up questions and one delayed question. Some topics were also extended by asking for advantages and disadvantages, although this was a skill targeted by the test and may have only been used due to the students' awareness that doing so would improve their grade in the class. This led to some participants taking an additional turn to add to the topic before moving on but only in restricted circumstances. There was also one topic where one participant only held the floor to ask a follow-up question and did not share their own ideas. This could be a sign of struggling to take the floor rather than a variation in turn distribution, however, this speaker did not have any issues taking the floor to ask a question and reacted to other speakers marking the floor as open with hesitation noises. The entire group also declined to add disadvantages when one speaker suggested it, indicating that they were instead comfortable changing the topic when there was nothing they wanted to say. These variations are closer to Sacks et al.'s (1974) description of organic conversational behavior, though they are still somewhat limited. From this data set it is not possible to determine whether this change is the result of increased comfort with floor management skills in the target language, increased comfort with floor management in videoconferencing interactions, or simply a minor variation without any relevant cause.

This behavior isn't necessarily surprising. Takao (1999) noted that Japanese university students participating in videoconferencing conversations tended to use monologue style turns and Tanaka and Mori (2022) found that in Japanese videoconferencing conversations, focus moved to content as opposed to interpersonal connection. Even in face-to-face interactions, Japanese learners of English have been observed to use repetitive turn orders (Young, 2015) and to have few instances of interacting with other speakers' ideas (Hennessy, 2020).

Further analysis of higher proficiency speakers of the same linguistic background and of groups with more variation in turn order and distribution could be valuable to understand more about the breadth of variations in behavior as well as the reasons for those variations.

Turn-taking techniques

Despite receiving similar instruction regarding techniques and phrases that could be used to transition between speakers, the three groups opted to use different techniques for floor management. Participants in discussion A used self-selection nearly exclusively and rarely marked the floor as open. The only exception is a single case of negotiation for meaning which resulted in six transitions between two participants where the next speaker was chosen either by name or by the content of the prior turn. In discussion B self-selection was common but there were occurrences of the current speaker selecting the next which tended to lead to smoother transitions with shorter gaps. Participants in discussion C, the most successful group, relied on self-selection for most turns, but frequently marked the end of speaking turns even when not selecting a next speaker.

Table 2 below shows a breakdown of turn-taking techniques used by the participants in each discussion. Table 3 shows the mean and median gap between turns when using each of these techniques. The included transitions are the same as those discussed in the previous section. Self-selection unmarked means that the current speaker finished their turn and did not do anything to specifically mark their turn as over and that the next turn was taken by a participant claiming the floor independently. Self-selection marked still means that the next speaker claimed the floor independently but that the current speaker did something to mark their turn as over, such as asking an opinion question without selecting who they wanted to answer. Other-selection means that the next speaker was specifically requested to take the floor by the current speaker. It is worth noting that discussion B has fewer transitions included as this group also had more transitions that were delayed due to marked thinking time or laughter. Again, the gap durations come from too small of a data selection to be valuable for quantitative analysis, but they are included here to help justify the selection of excerpts and commentary given in this section.

Table 2. Frequency of use of different turn-taking techniques

Number of occurrences	Self-selection unmarked	Self-selection marked	Other-selection	Total
Discussion A	10	3	6	19
Discussion B	7	3	6	16
Discussion C	8	13	10	31

Table 3. Duration of gaps (in seconds) according to turn-taking technique

Duration	Self-selection unmarked		Self-selection marked		Other-selection	
	Mean	Median	Mean	Median	Mean	Median
Discussion A	6.27	5.58	4.5	1.56	0.66	0.64
Discussion B	6.16	6.2	10.23	6.12	1.42	1.4
Discussion C	3.46	2.92	1.78	1.48	0.65	0.64

In discussion A the lack of verbal cues from the current speaker around turn transitions was very apparent. Most turns just ended when the current speaker stopped talking. The only examples of the current speaker selecting the next speaker were within an occurrence of negotiation for meaning. The only times the floor was marked as open were when introducing a new topic or when checking if the group had finished a topic. The heavy use of self-selection isn't inherently problematic and is a valid way to navigate speaker changes (Furo, 2001; Sacks et al., 1974), however it appeared inadequate for this group as the gaps between speakers appeared difficult to manage, as was discussed in the previous section. It is possible that using different techniques could have allowed this group to reduce the gaps between speakers.

Participants attempted to make use of body language to help with floor transitions with two of the three participants frequently looking at their screen as they finished a speaking turn. This may have been an attempt to use gaze to select the next speaker. However, research into videoconferencing has found that gaze and body language are less effective than in face-to-face interactions (Goodfellow et al., 1996; Imai et al., 2022; Reidl; 2021; Taylor, 2011) so this may not have helped. On the other hand, participants were capable of signaling that they wanted to hold the floor and not end their turn. Speakers regularly withheld the falling intonation that usually occurs at the end of a sentence (Ford & Thompson, 1996) to signal that they wished to continue speaking. The presence of techniques for holding the floor through CTRPs and the ineffective use of gaze and body language around turn transitions could indicate that these speakers had not fully adapted to videoconferencing interactions and had not yet identified methods that would help them in the medium.

The participants in discussion B made more use of turn-taking techniques with varied success. On a few occasions speakers used open ended questions to mark the end of their turns and there were six transitions that occurred when the current speaker selected the next speaker by name. The open-ended questions did not seem to help with transition speed, but of the six transitions where the next speaker was selected by name, five had a gap of less than two seconds and the sixth was longer due to unintentional overlap. The following example shows a successful transition using this technique.

Ryo: So if Japanese government raise taxes for rich people, they lost
Ryo: working motivation. /># So I think it is bad. />#

Ryo: How about you Saki? /># {1.08}

Saki: " Mmm, I think it is good. />#

However, there were occurrences where this technique seemed to be used due to a failure in self-selection. In four of the six occurrences there was a clear pause where the current speaker stopped speaking before resuming to choose the next speaker. One such pause was brief, but in the other three cases it seemed likely that the speaker decided to choose the next speaker because no one volunteered to take the floor. This next example best illustrates this point as the speaker pauses after completing their turn, asks an open-ended question to signal that they were finished, paused again briefly, and then finally chose the next speaker by name.

Saki: Japan's social security costs,[are enormous because of this system. />#

Ryo: [*' Yes.

Saki: So this– that point– that is a point we should discuss. /># (4.16)

Yui: [*' Yes.

Ryo: [*' Yeah.

Saki: Yeah. (1.88) What do you think about this view? /># (1.40)

Saki: Yui? /># {1.36}

Yui: " Oh, I um, I didn't this service. /># Saki said. />#

Thus, it seems that the group did identify a technique that helped them with floor transitions but had not yet developed the ability to use it consistently.

As in discussion A, the participants in discussion B often looked at their screens as they finished speaking. In this group, it seemed to occasionally be effective at marking the end of a turn as other participants often offered back-channels and supportive gestures in response, but it did not help to select the next speaker. This also occurred when using the technique of selecting the next speaker by name, with the current speaker looking at the screen as they finished talking but before asking another participant to speak. This could suggest that they were hoping to choose the next speaker with gaze but changed their strategy when they realized it was not successful.

Participants in discussion C primarily used self-selection to choose the next speaker. There were two occurrences of the current speaker selecting the next, but one of those was to select a speaker who had yielded the floor due to overlap at the previous floor transition. However, unlike in discussion A, speakers often marked the end of their turns with questions to help listeners identify when the floor was open. As in discussion B these questions sometimes included a pause between the speaker ending their turn and adding a question to mark the floor as open. The example below clearly shows this behavior as there is a clear gap after Ken finishes speaking with no movement towards a floor transition, but a quick transition to the next speaker once he resumes speaking to mark the end of his turn with a question.

Ken: So I think it's a good way for me. /># [(03.64)

Shun: [*' I think so. I think so.

Mai: [*' Mmhhh.
Yuna: [*' Yes.
Ken: Is there anything to add? /># {0.64}
Shun: " Yeah. Can I say? />#

This suggests that speakers were able to identify when floor transitions were not proceeding smoothly and use questions to help with the transition. However, unlike in discussion B they did not find it necessary to identify who should speak next. Additionally, many transitions happened quickly by self-selection without the need for any particular marking. The example below illustrates this well as Mai does nothing to mark the end of her turn, but Yuna self-selects as the next speaker fairly quickly.

Mai: I agree with Yuna and Shun. /># It is unfair, /># I think. /># {1:48}
Yuna: [*Mmhhh.
Shun [*Mmhhh.
Yuna: Can I say something? />#

The participants in discussion C did not attempt to use gaze as frequently or obviously as in other groups. Speakers did look at their screen at the end of their turns, but it was generally brief and indistinguishable from glances they made to their screen at other times. There was no reason to think they were trying to use body language for assistance with turn transitions.

Looking at the three discussions together we can see that different groups do manage their turns differently. Research into videoconferencing interactions has previously found that speakers do develop methods to assist with floor management (e.g. Moorhouse et al., 2022). This study suggests that effective strategies can vary between groups and some groups may not readily find a technique that can prevent frequent gaps in conversation. Participants in discussion C were very effective with self-selection while that was inadequate for discussion A. In discussion B selecting the next speaker by name improved transitions, but the technique seemed unnecessary for discussion C. It did seem clear that marking the end of a turn helped with transitions as participants in discussions B and C both benefitted from this technique, but whether it is valuable for the current speaker to select the next speaker or if self-selection can be adequate will depend on the participants in a given interaction.

Other observations: Backchannels and overlap

Two other factors that potentially contributed to the groups' turn-taking performances were the use of backchannels and avoidance of overlap. The group with the shortest gaps, discussion C, utilized high energy backchannels more frequently than other groups. Given the higher occurrence of backchannels in Japanese (Fujimoto, 2010; Furo, 2001; Kern, 2009) this could have helped with speaker comfort. This group was also the most comfortable with accidental overlap. Table 4 below shows the frequency of verbal backchannels used

during another speaker’s turn as well as the frequency of unintentional overlaps in each discussion.

Table 4. Backchannels and unintentional overlaps

	Discussion A	Discussion B	Discussion C
Backchannels while another person is speaking	25	67	207
Unintentional overlaps	0	2	9
Discussion duration (minutes)	12	16	16

In discussion A, support for the current speaker through backchannels was limited. In one case a speaker held the floor for 40.42 seconds and received a total of three backchannels, two of which were simultaneous. The backchannels that were present were also quiet and low energy. Listeners did nod their heads frequently, but since body language is less effective with videoconferencing (Reidl, 2021; Taylor, 2011) this would have been less valuable than a verbal reaction. In discussion B backchannels were more consistent, but similarly quiet and low energy. Listeners frequently reacted strongly after a speaker finished speaking, suggesting they were not comfortable reacting during another speaker’s turn. Participants in discussion B tended to use large gestures, perhaps as an attempt to overcome the reduced effectiveness of body language, but there was no evidence that this was successful. It is possible that the lack of backchannels indicated difficulty in identifying CTRPs, which could be one explanation for gaps at turn transitions. However, confirming this would require interviewing participants about their experience.

In discussion C, participants made more frequent use of backchannels during other speakers’ turns that also subjectively seemed to have more energy, potentially increasing their effectiveness at showing interest and engagement. Physical reactions were used, but they were usually combined with a simultaneous verbal backchannel. In addition to more closely matching Japanese conversation styles, this may also be beneficial for comfort in videoconferencing interactions, in particular as observed by Moorhouse et al. (2022).

The frequency of overlap and reaction of participants was also different between groups. Participants in discussion C appeared significantly more comfortable with unintentional overlap around turn transitions when compared to the other groups.

In discussion A there were no instances of overlap outside of backchannels and confirmations of understanding. The biggest overlap was two speakers simultaneously using short sentences to confirm they understood the current speaker as that speaker stopped talking. There were no examples of two speakers trying to take the floor at the same time and no participant ever tried to take the floor before the current speaker had finished. Since overlap was not observed it cannot be concluded that it would have been problematic,

but it may be connected to Kitamura's (2001) observation that L1 speakers of Japanese avoid overlap when communicating in English.

Discussion B had two examples of accidental overlap. Both were at turn transitions and easily solved, but in both cases the speaker who stopped talking showed embarrassment about creating the overlap. The participants in this discussion clearly seemed uncomfortable when accidental overlap did occur. In the example selected below two speakers tried to claim the floor simultaneously following a long pause and one, Saki, apologized multiple times as she yielded the floor.

Yui: So I think it's a good idea. /># {08.60}

Saki: *' Oh.

Ryo: *' Yes.

Saki: *' Okay. (5.88)

Saki: " Is [there any ahh, okay okay. Sorry. Sorry.

Ami: [" Can I say something? />#

Unintentional overlap in discussion C was more common with nine occurrences at turn transitions – three where a speaker began their turn before the current speaker finished and six where two speakers began to speak at the same time. All occurrences were brief, rapidly repaired, and unproblematic. In addition, the participants did not seem uncomfortable, and the discussion proceeded easily. In this example, we see a participant claim the floor through self-selection as another speaker is still finishing their turn.

Shun: So it's very good. /># [Good relation. />#

Mai: ["= Can I say something? />#

Ken: Okay.

Yuna: Okay. (0.84)

Mai: From the people who do volunteering choose volunteering they

Mai: can learn many things.

As the current speaker, Shun, appeared to be finished anyway, the rest of the group acknowledged Mai's claim to the floor, and she proceeded with her turn. No issues arose from the overlap and no commentary or reaction given to the overlap.

The six examples of two speakers starting simultaneously were similarly resolved without difficulty, with one speaker stopping and yielding the floor while the other continued to speak. The example shown here illustrates how easily the group was able to repair occurrences of overlap.

Shun: What do you think? /># {02.24}

Mai: " So, [Okay. I agree with Yuna. />#

Yuna: [" Can I? oh okay. /># Sorry. />#

Mai: Okay. I agree with Yuna and Shun.

When Yuna chose to yield the floor she apologized for the overlap, but the apology was brief and did not seem to cause discomfort for the participants and the discussion continued without any problems. Immediately following Mai's turn, the group allowed Yuna to claim the floor, suggesting that they were able to track this overlap and ensure that the speaker who yielded the floor was not skipped. The lack of difficulty around accidental overlaps, as is seen in the preceding example, made the interaction feel smoother and indicated that the group knew how to handle overlap and was not bothered by its existence.

When considering how the three groups behaved around overlap, it is possible to infer that comfort with overlap may be beneficial for turn-taking in videoconferencing conversations. Overlap has been observed to be problematic in videoconferencing interactions, often leading to stoppages and silence (Gettliffe & Erkat, 2016; Nakaya & Okamoto, 2022; Smidt et al., 2017). Kitamura (2001) also noted that Japanese speakers of English try to avoid overlap even in face-to-face interactions. This study suggests that overlap may be problematic for some speakers when interacting through videoconferencing software, but that speakers and groups that are more capable of resolving overlap may reduce its negative impact. The groups that had longer gaps between turns also seemed to be avoiding overlap whereas the group that appeared comfortable with overlap was much faster at transitioning between speakers.

Conclusion

While the analysis shows that the participants were able to handle turn-taking at a basic level in foreign language videoconferencing discussions, there is variability in the speed with which different groups were able to change speakers as well as in the techniques that were effective in managing these transitions. As was observed in prior research into videoconferencing interactions, silence and gaps between speakers could be considered problematic and the variation between groups indicated that the groups with longer gaps between speakers may have benefitted from using different techniques for floor management. The lack of interaction between speakers and tendency towards monologue turns was also a weak point of the interaction quality.

The issue of long gaps between speakers varied between groups with three specific observations. First, the use of a verbal technique to mark the end of a speaker's turn effectively reduced the gap between speakers. This technique was beneficial in both discussions where it was used. The variation between groups suggests that some groups may benefit from having the current speaker select the next speaker while other groups may find this unnecessary. Second, the increased prevalence and energy of backchannels observed in the discussion with the shortest gaps suggests that some form of support for speakers may be valuable. With videoconferencing software that features full-duplex audio, participants can react to a speaker without creating issues in comprehension and making use of the technology by providing enthusiastic support which may be beneficial for videoconferencing interactions. Finally, comfort with overlap and understanding of repair techniques may lead to an increased comfort with

claiming the floor. The two groups with a high frequency of long gaps strongly avoided overlap while the group that was able to transition quickly was more comfortable with accidental overlap. Learners for whom overlap is not inherently problematic may be able to better manage turns during videoconferencing interactions.

With regards to developing interaction between speakers within a single topic, fewer variations were observed. Follow-up questions and questions to extend a topic led to a small increase in the number of turns per topic and more in-depth discussion, suggesting an increased use of questions could assist with interaction. Videoconferencing discussions have been observed to feature monologue style turns and an increased focus on content at the cost of interaction (Takao, 1999; Tanaka & Mori, 2022) so this is somewhat expected, but the general avoidance of holding the floor multiple times for a single topic reduces the participants' ability to contribute. This could be a productive target of further research both in terms of identifying what behaviors are common in groups that are more successful and in identifying the beliefs of participants about the appropriateness of their own participation.

Looking ahead, evaluating how learners perform with mixed proficiency groups and with speakers who do not share a first language could give data into how these challenges can be addressed for non-classroom interactions. The techniques that were successful here may not work in different contexts and overly focusing on questions to extend a topic or on marking the end of a turn to aid in floor transitions would not necessarily be beneficial in different use cases. This study provides insight into how this particular demographic was able to negotiate the floor, but it is important to consider how learners will actually use the language and technology in the future.

References

- Barraja-Rohan, A. (2011). Using conversation analysis in the second language classroom to teach interactional competence. *Language Teaching Research*, 15(4), 479–507. <https://doi.org/10.1177/1362168811412878>
- Boland, J.E., Fonseca, P., Mermalstein, I., & Williamson, M. (2021). Zoom disrupts the rhythm of conversation. *Journal of Experimental Psychology: General*, 151(6), 1272–1282. <https://doi.org/10.1037/xge0001150>
- Celce-Murcia, M. (2007). Rethinking the role of communicative competence in language teaching. In E.A. Soler and M.P.S. Jorda (Eds.), *Intercultural language use and language learning* (pp 41–57). Springer.
- Clancy, P. M., Thompson, S. A., Suzuki, R., & Tao, H. (1996). The conversational use of reactive tokens in English, Japanese, and Mandarin. *Journal of Pragmatics*, 26(3), 355–387. [https://doi.org/10.1016/0378-2166\(95\)00036-4](https://doi.org/10.1016/0378-2166(95)00036-4)
- Cook, G. (1989). *Discourse*. Oxford: Oxford UP.
- Cui, X. (2014). Getting to the source: An instrument for examining the dynamics of problematic interactions. *RELJ Journal*, 45(2), 197–210. <https://doi.org/10.1177/0033688214535053>

- Dörnyei, Z., & Thurrell, S. (1994). Teaching conversational skills intensively: Course content and rationale. *ELT Journal*, 48(1), 40–49.
<https://doi.org/10.1093/elt/48.1.40>
- Du-Babcock, B. (1999). Topic management and turn taking in professional communication: First- versus second-language strategies. *Management Communication Quarterly*, 12(4), 544–574.
<https://doi.org/10.1177/0893318999124003>
- Du-Babcock, B. & Tanaka, H. (2013). A comparison of the communication behaviors of Hong Kong Chinese and Japanese business professionals in intracultural and intercultural decision-making meetings. *Journal of Business and Technical Communication*, 27(3), 263–287.
<https://doi.org/10.1177/1050651913479918>
- Ford, C., & Thompson, S.A. (1996). Interactional units in conversation: syntactic, intonational, and pragmatic resources for the management of turns. In E. Ochs, E.A. Schegloff, & S.A. Thompson (Eds.), *Interaction and grammar* (pp 134–184). Cambridge University Press.
- Fujimoto, D. (2010). Connecting EFL group discussions to research. *Bulletin of Osaka Jogakuin 2-year College*, 40, 1–17.
<http://ir-lib.wilmina.ac.jp/dspace/handle/10775/913>
- Furo, H. (2001). *Turn taking in English and Japanese: Projectability in grammar, intonation, and semantics*. Routledge.
- Gettliffe, N. & Erkat, O. (2016). Evaluating second language oral skills in an on-line videoconferencing environment. *8th international conference on education and new learning technologies proceedings*, 2838–2845. IATED Academy. <https://doi.org/10.21125/edulearn.2016.1617>
- Goodfellow, R., Jefferys, I., Miles, T., & Shirra, T. (1996). Face-to-face language learning at a distance? A study of a videoconference try-out. *ReCALL*, 8(2), 5–16. <https://doi.org/10.1017/S0958344000003530>
- Harumi, S. (2001). The use of silence by Japanese EFL learners. *PAC3 at JALT 2001 conference proceedings*, 27–34.
<https://jalt-publications.org/archive/proceedings/2001/027.pdf>
- Hennessy, J. (2020). Teaching journal: Reflections on developing interaction and improving turn-taking in high intermediate students. *New Directions in Teaching and Learning English Discussion*, 8, 23–29.
<https://doi.org/10.14992/00019170>
- Imai, R., Go, K., Tominaga, S., Kimura, Y., Sakai, T., Kojima, K., & Miyata, A. (2022). Effects of visualization of the flow of conversation on videoconferencing. 情報処理学会論文誌 (*Information Processing Society of Japan Journal*), 63(1), 11–20. <https://doi.org/10.20729/00215705>
- Jenkins, J. (2009). *World Englishes: A resource book for students*. Routledge.
- Kachru, B. B. (1990). *World Englishes and applied linguistics*. World Englishes, 9(1). 3–20. <https://doi.org/10.1111/j.1467-971X.1990.tb00683.x>
- Kasper, G. (1992). Pragmatic transfer. *Second Language Research*, 8(3), 203–231. <https://doi.org/10.1177/026765839200800303>

- Kawashima, H., Nishikawa, T., & Matsuyama, T. (2008). Visual filler: facilitating smooth turn-taking in video conferencing with transmission delay. *Proceedings of the 2008 conference on human factors in computing systems*, 3585–3590. <https://doi.org/10.1145/1358628.1358895>
- Kern, J.Y. (2009). To follow or to flout? Communicative competence and the rules of turn taking. *The Language Teacher*, 33(9), 3–9. <https://jalt-publications.org/tlt/articles/487-follow-or-flout-communicative-competence-and-rules-turn-taking>
- Kitamura, N. (2001). *Politeness phenomena and mild conflict in Japanese casual conversation*. [Doctoral Dissertation, The University of Sydney]. The University of Sydney Library. <https://ses.library.usyd.edu.au/handle/2123/844?mode=full>
- Lee, G. (2009). Speaking up: Six Korean students' oral participation in class discussions in US graduate seminars. *English for Specific Purposes*, 28(3), 142–156. <https://doi.org/10.1016/j.esp.2009.01.007>
- Marshall, D. T., Shannon, D. M., & Love, S. M. (2020). How teachers experienced the COVID-19 transition to remote instruction. *The Phi Delta Kappan*, 102(3), 46–50. <https://doi.org/10.1177/0031721720970702>
- Maynard, S. K. (1986). On back-channel behavior in Japanese and English casual conversation. *Linguistics* 24(6), 1079–1108. <https://doi.org/10.1515/ling.1986.24.6.1079>
- Moorhouse, B. L., Walsh, S., Li, Y., & Wong, L. C. (2022). Assisting and mediating interaction during synchronous online language lessons: Teacher's professional practices. *TESOL Quarterly*, 56(3), 934–960. <https://doi.org/10.1002/tesq.3144>
- Morita, N. (2000). Discourse socialization through oral classroom activities in a TESL graduate program. *TESOL Quarterly*, 34(2), 279–310. <https://doi.org/10.2307/3587953>
- Nakaya, A. & Okamoto, M. (2022). Overlapping talk in virtual meetings: Focusing on interference with preceding utterances. *Proceedings of the Japanese society for artificial intelligence: JSAI technical report, SIG-SLUD*, 94, 70–77. https://doi.org/10.11517/jsaislud.94.0_13
- Nikou, S. A. (2021). Web-based videoconferencing for teaching online: continuance intention to use in the post-COVID-19 period. *Interaction Design and Architecture*, 47, 123–143. <https://doi.org/10.55612/s-5002-047-006>
- Riedl, R. (2022). On the stress potential of videoconferencing: definition and root causes of Zoom fatigue. *Electronic Markets*, 32, 153–177. <https://doi.org/10.1007/s12525-021-00501-3>
- Ryan, J. & Forrest, L. (2021). 'No chance to speak': developing a pedagogical response to turn-taking problems. *Innovation in Language Learning and Teaching*, 15(2), 103–116. <https://doi.org/10.1080/17501229.2019.1687709>
- Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, 50(4), 696–735. <https://doi.org/10.2307/412243>

- Saito, H. & Beecken, M. (1997). An approach to instruction of pragmatic aspects: Implications of pragmatic transfer by American learners of Japanese. *The Modern Language Journal*, 81(3), 363–377.
<https://doi.org/10.1111/j.1540-4781.1997.tb05497.x>
- Sierra, J. (1999). Real linguistic experiences using chat sessions or videoconferencing. *The Internet TESL Journal*, 5(3).
<https://eric.ed.gov/?id=ED427526>
- Smidt, E., McAndrew, A., & McDyre, B. (2017). Synchronous video chat sessions in a TESOL online graduate course: Instructor roles and best practices. In I. Sanz-Sanchez, S. V. Rivera-Mills, & R. Morin (Eds.), *Online language teaching research: Pedagogical, academic and institutional issues* (pp 79–99). Trysting Tree Books.
- Sowden, C. (2012). ELF on a mushroom: the overnight growth in English as a Lingua Franca. *ELT Journal*, 66(1), 89–96. <https://doi.org/10.1093/elt/ccr024>
- Stivers, T., Enfield, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., Hoymann, G. Rossano, F., de Ruiter, J. P., Yoon, K. E., & Levinson, S. C. (2009). Universals and cultural variation in turn-taking in conversation. *PNAS*, 106(26), 10587–10592. <https://doi.org/10.1073/pnas.0903616106>
- Takahashi, T. & Beebe, L. M. (1987). The development of pragmatic competence by Japanese learners of English. *JALT Journal*, 8(2), 131–155.
<https://jalt-publications.org/jj/articles/2838-development-pragmatic-competence-japanese-learners-english>
- Takao, S. (1999). The effects of narrow-band width multipoint videoconferencing on group decision making and turn distribution. *ACM SIGSOFT Software Engineering Notes*, 24(2), 109–116.
<https://doi.org/10.1145/295666.295678>
- Tanaka, H. (1999). *Turn-taking in Japanese conversation: A study in grammar and interaction*. John Benjamin's Publishing Company.
- Tanaka, H. (2006). Emerging English-speaking business discourses in Japan. *Journal of Asian Pacific Communication*, 16(1), 25–50.
<https://doi.org/10.1075/japc.16.1.03tan>
- Tanaka, H. (2008). Communication strategies and cultural assumptions: an analysis of French-Japanese business meetings. In S. Tietze (Ed.), *International management and language* (pp 154–170). Routledge.
- Tanaka, S. & Mori, N. (2022). Qualitative differences between online and face-to-face conversation from the viewpoint of intercorporeality. *Journal of Epistemology and Mind Sciences*, 4(1), 2–17.
https://doi.org/10.50882/epstemindsci.4.1_2
- Taylor, T. (2011). Video conferencing vs talking face-to-face: Is video suitable for supportive dialogue? *International Journal of Therapy and Rehabilitation*, 18(7), 392–402. <https://doi.org/10.12968/ijtr.2011.18.7.392>
- Tecador Cabrero, M. (2013). *Developing interactional competence through video-based computer-mediated conversations: Beginning learners of Spanish*. (Publication No. 3595159) [Doctoral dissertation, The University of Iowa]. ProQuest Dissertation Publishing.
<http://doi.org/10.17077/etd.t6sm6wlt>

- ten Have, P. (2007). *Doing conversation analysis* (2nd Edition). SAGE Publications. <https://doi.org/10.4135/9781849208895>
- van Os, M., de Jong, N., H., & Bosker, H. R. (2020). Fluency in dialogue: Turn-taking behavior shapes perceived fluency in native and nonnative speech. *Language Learning*, 70(4), 1183–1217. <https://doi.org/10.1111/lang.12416>
- Williamson, J. (2019). A pragmatics explanation for Japanese-English turn-taking contrasts and the need for pedagogical intervention: A response to Davey Young's TLT article. *The Language Teacher*, 43(1), 14–18. <https://doi.org/10.37546/JALTTLT43.1-3>
- Wilson, T. P. & Zimmerman, D. H. (1986). The structure of silence between turns in two-party conversation. *Discourse Processes*, 9(4), 375–390. <https://doi.org/10.1080/01638538609544649>
- Wong, J., & Waring, H. Z. (2010). *Conversation analysis and second language pedagogy*. Routledge.
- Young, D. (2015). A conversation analysis of the acquisition and use of turn-taking practices in an English discussion class. *New Directions in Teaching and Learning English Discussion*, 3, 320–330. <https://doi.org/10.14992/00015992>
- Young, D. (2018). Contrastive models for turn-taking in English and Japanese. *The Language Teacher*, 42(3), 9–12. <https://doi.org/10.37546/JALTTLT42.3-2>

Appendix A

Transcription key and analysis symbols

- [overlapping utterances
- () noticeable pause
- { } time between floors (disregarding backchannels)
- cut-off sound
- = latched utterance
- ? rising intonation
- . falling intonation
- , continuing intonation
- / grammatical completion point
- > semantic completion point
- * backchannel
- # CTRP (Complex Transition Relevance Place)
- " floor taking speaker change
- ' non-floor taking speaker change
- :: non-linguistic action