LingBaW *Linguistics Beyond And Within 9 (2023), 36-48*

Representing temporal concepts using redundant gestures in L2 ongoing interactions

Hiroki Hanamoto
Tokyo Denki University, Japan

Abstract
Human conversational interaction is multimodal, involving both verbal and non-verbal modalities. That is, when a speaker and listener interact, they use not only spoken messages but also manual gestures. Manual gestures and spoken messages are semantically and temporally related and work together to create and express a complete meaning. This study employs a data-driven approach to investigate how L2 learners spontaneously employ gestures to express temporal concepts in ongoing dyadic interactions using 11 recorded interactions among L2 learners. The distribution and frequency of specific types of manual gestures were examined using sequential and gesture analyses. The results showed that, when representing temporal concepts, the participants produced language-redundant gestures. For example, to convey temporal concepts, they tended to co-express the same information with manual gestures, namely abstract deictic and metaphoric gestures, on an imaginary mental timeline axis, which appeared to represent the English grammatical concepts of tense and aspectual meaning. Regarding the functional differences in gestures in interactions, based on sequential analysis, speakers employed language-redundant gestures to express time concepts explicitly in comprehension sequences and in the negotiation of meaning as a strategy of repair for lexical retrieval, paraphrasing, and clarification. These findings reveal that understanding the use of both modalities, speech and gestures, is critical in uncovering how speakers conceptualize time in their minds and integrate space and time in language.

Keywords: Redundant gesture, multimodality, temporal concept, L2 interaction

1. Introduction

Researchers have long acknowledged the co-development of speech and gestures, especially in speech production (McNeill 1992). Given that people invariably gesture when they speak, these co-speech gestures are commonly assumed to be related to speech and are often coordinated at a semantic and rhythmic level (Kendon 1980; McNeill 1992). McNeill's Growth Point Theory (McNeill 1992) assumes that speech and gestures merge and that both modalities become an integrated system in spontaneous talk. Whereas gesture-speech production is mandatory in language comprehension, the relationship between gesture and speech is considered to be associated with different types of gestures, and the meaning of gestures depends on the accompanying speech (Kendon 2004). Gestures have speaker-oriented and listener-oriented

functions, implying that gestures fulfill a range of functions in spoken language (Goldin-Meadow 2000). If people gesture during spoken interactions, they might produce co-speech gestures for both internal cognitive and communicative purposes (Kita 2000). Accordingly, the author investigated the use of co-speech gestures when talking and representing time concepts in ongoing interactions. Thus, this study examines the use of temporal gestures in face-to-face dyadic interactions and aims to understand how the dyads conceptualize time and integrate space and time in language. Furthermore, this study reveals that understanding the use of both modalities, speech and gestures, is critical in social interaction and language-learning contexts. Gestures may be considered visual aids that efficiently provide cues, transforming an abstract concept into something concrete and visible for effective communication.

2. Literature review

2.1. Gestures and speech

Several studies, including those by Kita (2000) and McNeill (1992), have reported that speakers unconsciously employ gestures during speech. The exact role of gestures remains unclear. However, a substantial body of gesture studies offers two different views: Gestures have beneficial effects on addressers (e.g., Kita et al. 2007; Krauss and Hadar 1999) and on addressees (e.g., Gullberg 1998; Kendon 1994). Pinpointing the precise functions of gestures is not easy because past studies have employed varied methods and approaches. However, a holistic view suggests that gestures have both cognitive and communicative functions.

Speakers employ spontaneous manual movements that accompany speech, and their meaning relies heavily on verbal expression. According to Sime (2006), spontaneous gestures are defined as mostly unconscious speech-related manual movements that are directly related to the local speech context. Therefore, spontaneous gestures only occur during speech and are synchronized with it rhythmically, semantically, and pragmatically (McNeill 1992). Many scholars have reviewed the relationship between speech and gestures, and several have adopted a clarification system introduced by McNeill (1992). According to McNeill's taxonomy, spontaneous gestures are of four types: iconic (reflecting the property of the referent or action), metaphoric (referring to an abstract notion), deictic or pointing (indicating objects or events in both a concrete and an abstract sense), and beats (adding emphasis or rhythmic coordination in speech). That is, iconic, metaphoric, and deictic gestures, referred to as representational gestures, are all semantically related to their accompanying speech.

The role of gestures in the expression of semantic coordination of speech has been investigated in many ways. However, whether gestures refer to a redundant entity or to accompanying speech, or whether they supplement speech, has been debated (e.g., Hostetter 2011). In other words, gestures can depict no more information that was not already conveyed verbally (redundant gestures) or gestures include additional or supplemental information not present in the verbal element (non-redundant gestures). People can communicate through language as well as other semiotic modes, including gestures. Therefore, it would make more sense that gestures complement the speaker's message with what they are willing to express

verbally as visible actions (Kendon 2004). Meanwhile, why speakers sometimes employ redundant gestures when speech itself already provides sufficient information is still under discussion, and many answers and differing views have been suggested, including visual understandability (e.g., Tuite 1993), semantic integration of gesture and speech (e.g., Kita and Özyürek 2003), or preempting problems of ambiguity and vagueness (e.g., de Ruiter et al. 2012).

2.2. Gestures for expressing temporal concepts

When people think about time, they usually convey spatial and temporal concepts through both speech and gestures (Casasanto and Jasmin 2012), and, according to Lakoff and Johnson (1980), sometimes map time onto imaginary metaphorical axes in space. Space is exploited to handle temporal concepts visually, and according to Kendon (1990), abstract deictic and metaphoric gestures along time axes that accompany speech can be used to represent time concepts. For example, Torralbo et al. (2006) pointed out that English and Spanish speakers tend to conceptualize mental timelines laterally, from left to right, and along the sagittal plane, from back to front. Specifically, moving our hands to the left or back signifies the past along the timeline, whereas movements to the right or front represent the future. Thus, space mapping with speech is one option for handling temporal concepts, namely past and future timeline concepts.

Here, a question that might emerge is whether people gesture the same way or share the same mental timeline when talking about temporal concepts. It has been widely reported that time is spatially conceptualized and activated in a speaker's mind and is mapped onto imaginary mental timeline axes. In any language, people recognize time conception along the sagittal axis: past events lie at the back, and future or coming events in front. However, directionality in space-time mapping varies by situation or context in interactions as well as according to the language-speaking culture. Casasanto and Jasmin (2012) claimed that the direction of reading and writing in language or the flow of time on calendars can reflect a speaker's mind to shape their mental timeline flow. Meanwhile, according to Cienki (1998), speakers tend to gesticulate laterally to avoid intruding upon interlocutors' physical personal space.

As noted above, researchers have been concerned with studies exploring gesture use to express time concepts. However, to the best of the author's knowledge, as compared to interactions between L1 and L2 interlocutors, scant research has examined L2 learners' use of temporal gestures in time conceptualization in interactions between L2 learners only. Thus, a lack of direct evidence exists that investigates the gestures used when representing time concepts among non-native speakers of English. The few empirical studies on the use of temporal gestures in L2 contexts suggest that L2 learners employ temporal gestures to indicate temporal concepts as well as interactional communication strategies. For instance, Hanamoto (2020, 2021) analyzed face-to-face dyadic interactions using an emic approach and qualitatively explicated the ongoing process of representing temporal concepts using multimodal sequential analysis. These case studies concluded that many temporal references from both modalities were displayed and that the participants compared the tense timeline using gestures such as the abstract deictic and metaphoric types. Alibali et al. (2013) found similar patterns of abstract deictic gestures in highlighting grammatical tense concepts as interactional resources. Hanamoto (2020, 2021) further reported that participants used temporal gestures as

supplementary visual information or to replace spoken output. These results suggest that gestures accompanying speech indicate temporal concepts as well as highlight the communicative relevance of the information expressed by the gesture.

These studies raise simple but interesting questions. Looking over the studies of temporal gestures accompanying speech in interactions, there is still little evidence from naturally occurring data, not using a controlled or task-oriented conversation, as compared with lexical production or explanation. In addition, studies exploring how L2 learners themselves co-create the conceptualization of time visually and auditorily are yet to be systematically studied. Therefore, the present study followed a data-driven approach to broaden the understanding of gesture use in the ongoing process of interaction. Specifically, how and for what purpose do L2 learners visually co-create the conceptualization of time in ongoing interactions? The present study thus contributes to gesture studies, in particular with regard to interactional multimodal strategies in social interaction and language learning contexts.

3. Method

3.1. Participants

The participants were 18 male and 4 female undergraduate students studying science and engineering in Japan. They were all non-native English speakers and had learned English as a second language. Their L1 backgrounds were Japanese (n = 11), Cantonese (n = 3), Arabic (n = 2), Vietnamese (n = 2), Indonesian (n = 1), Nepalese (n = 1), Pekinese (n =1), and Turkmen (n = 1). According to the TOEIC-based placement test conducted at their university, their English levels ranged from beginner to intermediate. All participants voluntarily participated after reading and signing a consent form prior to data collection.

3.2. Data collection

Data were collected in the laboratory where the author worked. The author set up 11 dyads between Japanese and international students for video-recording interactions but did not control for their English proficiency levels. Nobody had met their dyad partner before the recording. The author set up two high-performance digital cameras with additional microphones connected to each speaker's clothes. The two cameras simultaneously and clearly recorded the speaker's speech and manual gestures. The recording time for the interactions was approximately 15 minutes. No specific discussion topics were assigned, and the participants were asked to start a free conversation. The type and purpose of the interaction could influence or relate to the communication strategies used. Thus, before the interaction, the author asked each dyad to decide on the first topic to talk about and to manage, create, and develop ongoing interactions between the pairs. The result showed that they preferred "safe topics" (Meierkord 2000) such as their common interests, majors, cultural differences, or hobbies. The interlocutors attempted to complete one topic in around 10 turns and moved on to the next topic. Given this structure, findings from non-institutional casual interaction settings provided insights into

interactional resources in L2 or English as a lingua franca (ELF) contexts. Accordingly, this study sought to determine how L2 learners represent temporal concepts through the use of spontaneous gestures in ongoing interaction.

3.3. Coding and data analysis

The author analyzed fragments in which the topic was related to the temporal concepts or references. First, based on the analytical framework of sequential analysis focusing on talk-on-interaction (Schegloff 1992), the author divided all 11 video-recorded interaction data items into fragments of sequential structure. Fragments for further analysis were transcribed and annotated using ELAN, an open-source multimodal annotation software. All utterances, including nonverbal behaviors, were transcribed verbatim using Jefferson's (1984) and McNeill's (2005) transcription conventions (see Appendix for transcription conventions). McNeill's typology was used to classify and code all spontaneous manual gestures into iconic, deictic, metaphoric, and beats gestures. Moreover, the gesture stroke and spatial direction may correlate to the timeline concepts implied by co-verbal intentions (e.g., Casasanto and Jasmin 2012). Gesture strokes, which were considered to convey the most significant and meaningful parts of the conversation, were also coded for viewpoints based on McNeill (1992) as follows: hand (left, right, bimanual), orientation (lateral, vertical, sagittal), and spatial direction (leftward, rightward, upward, downward, forward, center, backward, upper left, upper right). Gestural coding and identification were performed by two coders to ensure reliability. The substantial inter-rater agreement (κ = . 78) was calculated using Cohen's kappa coefficient.

4. Results and discussion

4.1. Temporal co-speech gestures along imaginary mental timeline axis

The author examined excerpts in which the participants used co-speech gestures when discussing time concepts. In other words, the author attempted to determine how and for what reason the participants co-created the conceptualization of time visually in ongoing interactions.

In the 165 minutes of talk, the author observed 67 cases of talking about timeline concepts; however, 37 of the excerpts were excluded from further analysis because these references were only conveyed by verbal expression, not visually. Thus, for further analysis, 30 spontaneous gestures with speech expressing temporal concepts or references were considered. These 30 utterances were expressed by both modalities, verbal and gestural. While the participants discussed temporal concepts through verbal utterances, they also expressed the same information visually with their hands. Therefore, by observing this match of speech and gesture content, redundant gestures would be more beneficial in expressing temporal concepts than would non-redundant gestures among L2 learners ranging from the beginner to the intermediate proficiency level.

Of the 30 temporal gestures that they produced along with the utterances, the gestures expressing temporal concepts were mostly abstract deictic and in some cases metaphoric

(Gullberg 1998): 21 utterances were accompanied by abstract deictic gestures, and the remaining were metaphoric gestures. As expected, there were no uses of iconic or beats gestures found in representing temporal concepts. The current investigation therefore focuses on deictic and metaphoric gestures.

All of the gestures were oriented along an axis with a clear direction: 22 gestures laterally and 8 sagittally (Figure 1). Speech and gestures are intertwined in communication; thus, co-speech gestures on both axes appear to function interactionally.

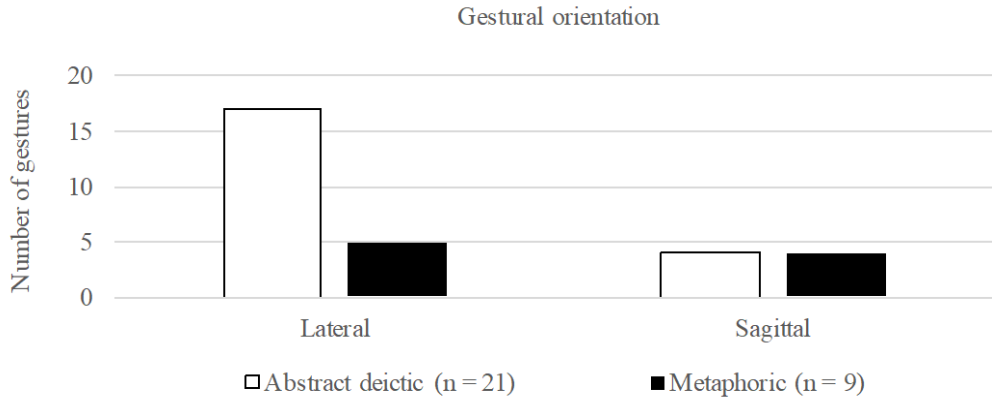

Figure 1: *Gestural orientation of temporal co-speech gestures*

The two graphs in Figure 2 show the distribution of the spatial direction of the imaginary timeline axes expressing earlier (past) and later (future) time. As this distribution of axis orientations shows, the participants placed and moved their hands onto the imaginary mental time axis when referring to timeline concepts (Lakoff and Johnson 1980). However, it is important to note that speakers in this study mapped their hands more on the lateral timeline than on the sagittal timeline axis when describing time concepts. That is, although they could employ more than one metaphorical time axis when referring to time concepts, they tended to arrange time from left to right (laterally) rather than from back to front (sagittally). This finding has been demonstrated in previous studies, including Casasanto and Jasmin (2012). Thus, as shown in Figure 3, in the case of referring to the earlier time, or the past tense or past events, they tended to move or point their hands into the space at their far left on a lateral imaginary timeline more than at their back along a sagittal axis. Meanwhile, regarding later time, they mainly gestured to the right rather than the front.

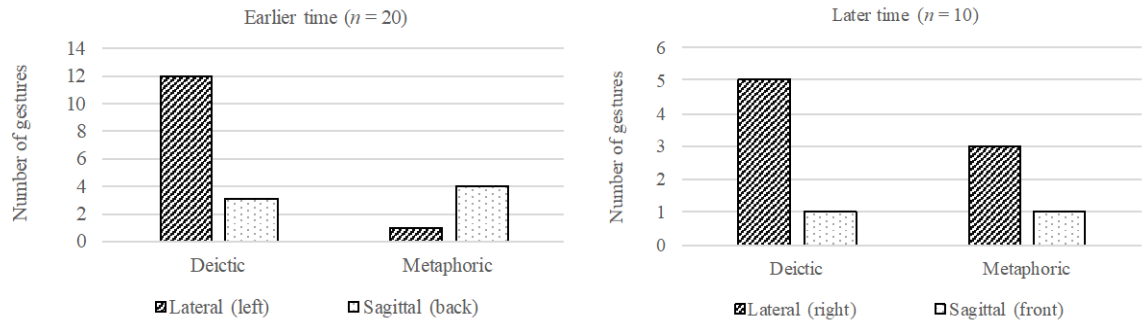

Figure 2: *Distribution of spatial direction of the imaginary mental timeline axes expressing earlier and later time*

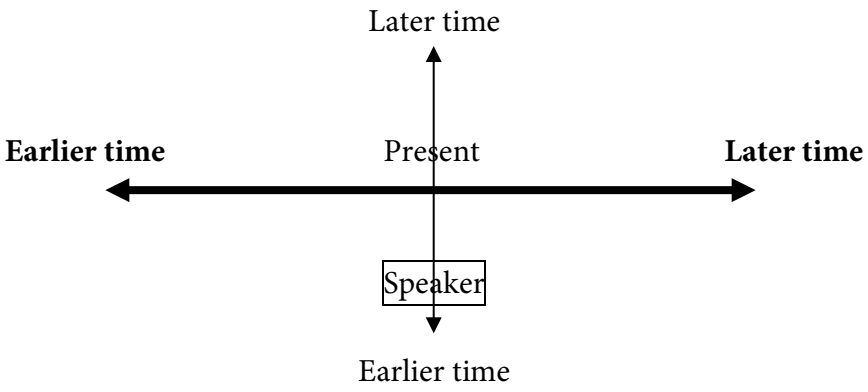

Figure 3: *Timeline axes for expressing tense references*

Each gesture on the imaginary temporal axis also appeared to adequately express different time concepts. For example, speakers tended to co-employ abstract deictic gestures along the lateral axis to accompany simple tense. Specifically, the participants extended their hands to the left to accompany simple past tense and to the right for simple future tense. Meanwhile, metaphorical gestures were mapped on the sagittal axis and synchronized with speech representing the grammatical concept of the aspect in English (i.e., present progressive and present perfect).

4.2. The function of the temporal co-speech gestures

The second goal of this study was to determine the function of the temporal co-speech gestures along the imaginary timeline axis. Combining traditional video annotation with sequential analysis, which addresses how talk construction is designed and negotiated, opens up new avenues to consider in analyses of the functions of temporal co-speech gestures. According to sequential analysis, the participants in this study expressed the same temporal concepts or references verbally as well as through language-redundant gestures along the imaginary mental timeline axis from different viewpoints, namely drawing on the concept of tense and aspectual meaning.

The author followed the transcriptions and video-recorded data with respect to functional differences in co-speech gestures. Based on sequential analysis, the author found that the participants employed language-redundant gestures when expressing time concepts, both in comprehension sequences where a speaker and a listener do not have outstanding problems in understanding and also in the negotiation of meaning, which shows a process of how participants in their interaction overcome communication problems in repair work. As Figure 4 shows, each use of gesture could be divided into four functions. Moreover, they were distributed in both situations, namely in the comprehension sequences (17) and repair negotiation sequences (13). More specifically, gestures employed in comprehension sequences were coded as "emphasis on a part of a reference" and the remaining three functions, "lexical retrieval," "paraphrasing," and "clarification," were observed in repair negotiation sequences. It is not possible to quantitatively compare the number of gestures between the two types because of the different total numbers of gestures; rather, we can compare the proportion of each function between the two types of gestures. Figure 4 shows that the participants gestured to emphasize a part of a time reference, which tends to be seen in the case of comprehension

sequences. These gestures did not contain information that gave additional meaning to the speech or generate explicitness of their verbal explanation or complements to interlocutors in interactions, but instead, speakers employed gestures that functioned as language-redundant gestures. Interestingly, information expressed by both modalities was explicitly conveyed to an interlocutor, which means such instances occurred during smooth and normal turn-taking, not repair sequences. In other words, gestures along imaginary mental timeline axes presented similar images through the dual mode, thus appearing to add redundancy to the semantics of speech that they accompanied to emphasize a part of the timeline reference.

Meanwhile, language-redundant gestures co-occurring with speech were also used in the sequences of negotiation of meaning as three types of strategies of repair: lexical retrieval, paraphrasing, and clarification. Here, it is noteworthy that such gestures were performed in connection with repair of prior talk by a speaker or a listener, namely same-speaker repair or other-speaker/recipient repair. For instance, in terms of lexical retrieval, speakers in interactions gradually moved and placed their hands on the axis to ease access to the timeline referent's word form production. In contrast, as an other-speaker repair, recipients also moved their hands to the axis with speech to increase the chance of being heard and understood to make an utterance intelligible to an interlocutor as a paraphrasing strategy; alternatively, for clarification and confirmation for resolving problems, recipients gestured alongside speech to promote the comprehension of a prior speaker's turn, which is interpreted as an interactional communicative strategy. Therefore, in case the construction and negotiation of meaning end up as incomplete, complementary combinations of gesture with speech are assumed to be effective in speech production and ongoing interactions. Excerpts 1–3 below illustrate how redundant gestures are used as strategies of repair, namely lexical retrieval, paraphrasing, and clarification, employed in their turn-taking.

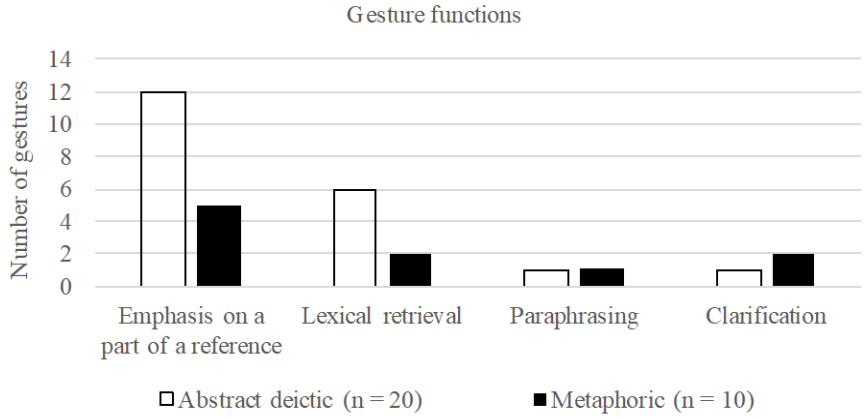

Figure 4: *Distribution of the abstract deictic and metaphorical gesture functions*

(4) Excerpt 1: Lexical retrieval

```
     A (m) = Vietnamese; B (m) = Japanese

  1.A    before I hmm go to university {I lived in}: {Tokyo}=
 →2.                                  {((moves his right fisted hand backward))}
 →3.                                                {((makes a thumbs-up and
 →4.     points a backward and sagittally twice))}
  5.B    =Tokyo
  6.A    ((nodding twice))
```

(5) Excerpt 2: Paraphrasing

```
     C (f) = Japanese; D (m) = Turkmen

  1.C   Japanese: Rapunzel? ((tilts her head to the left to appeal D to confirm
  2.    what she wanted to say))
  3.D   Rapunzel {yes=}
  4.             {((nodding twice))}
  5.C   =Rapunzel hmm. wear::: CHANGE? I go to: Tokyo Disney land
  6.D   (.) you {dress?=}
  7.           {((pretends to dress upper body bimanually))}
  8.C   =YES YES YES
  9.D   dress like (.) hmm Rapunzel?=
 10.C   =YES YES YES
→11.D   then you {went to a: Disney land?}
→12.             {((moves his right hand and points leftward))}
 13.C   {YES YES YES
 14.    {((nodding twice))}
```

(6) Expert 3: Clarification

```
     E (m) = Saudi Arabian; F (m) = Japanese

  1.E    so: you wanna speak like::any: English guy?
  2.F    hmm: no no ((shaking his head twice))
  3.E    [but:]
  4.F    [it's] my first time
 →5.E    it's first time {for you} but: your: any {your future
 →6.                    {((pointing to F with his right hand))}
 →7.                                            {((moves his right hand
 →8.     rightward))}
  9.F    maybe I need: yes
```

Additionally, the timing of the gestures further supports this finding. Figure 5 presents the timing between the onset of a lexical affiliate related to temporal concepts and the onset of a gesture. It is common for the timing of gestures and speech to be related to the familiarity of lexical features, and the gaps between speech-gesture onsets are related to difficulties with word retrieval (e.g., Krauss et al. 2000; Morrel-Samuels and Krauss 1992). According to this claim, a gesture's synchrony with speech or occurrence slightly before lexical affiliates implies the same or similar semantic information. Meanwhile, regarding speech-gesture asynchrony, where the gesture onset follows lexical affiliates, it might be possible to treat it as the perception of some problematic semantic references. In this study, we can see a similar tendency—that gestures functioned to emphasize a part of a reference and that clarification tended to precede or synchronize with the lexical affiliate. This means that speakers already know or familiarize themselves with these lexical affiliates and that they are more ready to facilitate speech production.

Understandably, these gestures were mainly observed in comprehension sequences, which means smooth turn-taking without encountering problems that needed repair when emphasizing a part of a reference with speech. Furthermore, the author found another set of gestures synchronized with speech in the repair sequences, as shown in Figure 5. However, they were used in the course of negotiation of meaning through repair strategies such as paraphrasing and clarification; they were not self-initiated but addressed and repaired by a recipient in the ongoing interaction. Such initiation of other-repair typically occurs in the turn immediately following a problematic display (Schegloff 1992). Thus, it can be interpreted that

the participants' non-verbal actions were synchronized with verbal speech to raise explicitness partly related to the repair of incomprehension.

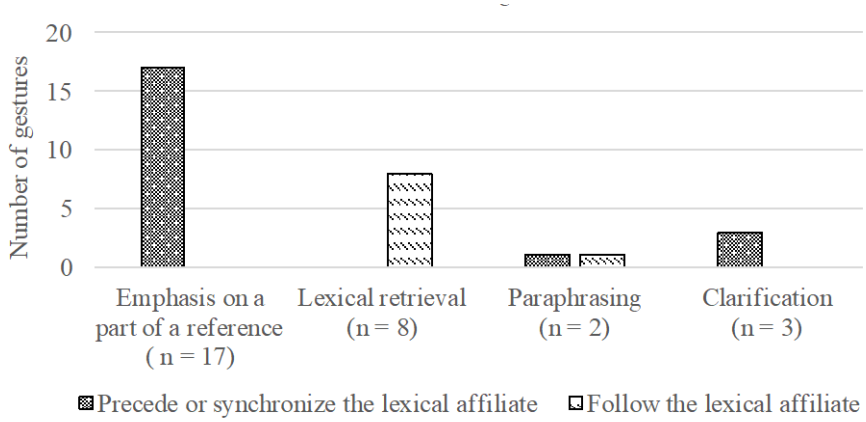

Figure 5: *Timing of gestures in relation to the temporal lexical affiliate in accompanying speech*

5. Conclusions

The goal of the present study was to describe how L2 learners spontaneously employ gestures when expressing temporal concepts in ongoing dyadic interactions. By using multimodal sequential analysis, the author obtained and explored 30 co-speech gestures for expressing temporal concepts.

The results showed that redundant gestures would be beneficial both in comprehension and repair negotiation sequences when representing temporal concepts. Specifically, the participants produced co-speech abstract deictic and metaphoric gestures when expressing temporal concepts. In other words, the visual information was largely redundant with speech (Wagner et al. 2014). This finding provides further support for the use of abstract deictic and metaphoric gestures with speech in terms of temporal expressions, as reported by Gullberg (1998). Gesturing with speech involves placing and moving hands along the imaginary mental time axis, more often on the lateral than on the sagittal axis (Casasanto and Jasmin 2012). Moreover, these spatial directions through manual gestures tended to represent the English grammar time concepts of tense and aspect. In other words, the present study's data show that accompanying gestures can reflect a speaker's mental representation of their conception of time. Additionally, the author explored the functions of language-redundant gestures in the expression of time concepts with speech. According to the sequential analysis and the timing of gestures and speech, the author found that the speakers employed gestures when expressing time concepts both in comprehension sequences and in the negotiation of meaning in repair work. That is, gestures performed in interaction fulfilled a range of functions, as Goldin-Meadow (2000) suggested, and confirmed in this study as an emphasis on a part of a reference or a repair strategy in the form of lexical retrieval, paraphrasing, and clarification, in the interactions of L2 learner ranging from the beginner to the intermediate proficiency level.

As this study provided additional evidence to support the claim that the participants tended to place and move their hands onto the imaginary mental time axis in the effective use of redundant gestures when referring to timeline concepts, it presents some methodological

weaknesses. The first limitation is related to the direction in which time is conceptualized spatially along people's mental imaginary timelines. As reported in the literature, including Casasanto and Jasmin (2012) and Lakoff and Johnson (1980), cross-cultural differences in the conceptualizations of time appear to vary according to speakers' L1 backgrounds or culture. While this study only considers the use of temporal co-speech gestures in ongoing dyadic interactions from different L1 backgrounds, interactions including pairs of participants from the same L1 background or culture also need to be examined in a cross-cultural context. A further limitation in this study was insufficient control in L2 proficiency level. The participants' English levels ranged from beginner to intermediate, which resulted in a lack of clear variability in a wide range of L2 proficiency level. Some experiments have shown that different levels of L2 proficiency may influence gesture use or gestural function including in a narrative (e.g., So et al. 2013) or role-playing context (e.g., Gregersen et at. 2009). Thus, the proficiency difference use of gesture with regard to redundant and non-redundant gestures, especially in L2 interactions, should be further investigated. Similarly, another limitation—disregarding the necessity for larger sample sizes for more qualitative and quantitative data analysis because people do not always employ gestures with speech, and ignoring individual differences—concerns the effect of mutual visibility or gaze on the number of gestures in face-to-face interaction, which a number of studies have examined (e.g., Bavelas et al. 2008; Enfield et al. 2007). Therefore, further research conducted by manipulating the visibility of the addresser and addressee is required to uncover the use of temporal co-speech gestures more accurately in ongoing interactions.

To conclude, as a direction for further research to overcome these limitations, it would be useful to determine the exact role of gestures in interactions from various domains, including naturalistic conversations described in the present study. That is, understanding the use of both modalities, speech and gestures, when expressing temporal references is crucial for uncovering how speakers conceptualize time in their minds and integrate space and time in language.

This is also a growing area of research, as understanding the relationship between language learning and gestures is important for nonverbal communication. Language or spoken messages are not always intelligible but rather pragmatically ambiguous. Gestures are visual aids that efficiently provide cues, transforming an abstract concept into something concrete and visible for effective communication. The presence of manual movements might facilitate language production and improve comprehension. This possibility segues into a discussion on whether speakers actually employ gestures for themselves (e.g., Kita and Özyürek 2003; Krauss and Hadar 1999) or intentionally for listeners (e.g., Gullberg 1998; Kendon, 2004) as well as the relationship between redundant and non-redundant gestures.

References

Alibali, M. W., Young, A. G., Crooks, N. M., Yeo, A., Wolfgram, M. S., Ledesma, I. M., Nathan, M. J., Church, R. B., and E. J. Knuth. 2013. Students learn more when their teacher has learned to gesture effectively. *Gesture* 13(2*)*: 210–233. https://doi.org.10.1075/gest.13.2.05ali

Bavelas, J., Gerwing, J., Sutton, C., and D. Prevost. 2008. Gesturing on the telephone: Independent effects of dialogue and visibility. *Journal of Memory and Language* 58(2*)*: 495–520. https://doi.org.10.1016/j.jml.2007.02.004

Casasanto, D., and K. Jasmin. 2012. The hands of time: Temporal gestures in English speakers. *Cognitive Linguistics* 23(4): 643–674. https://doi.org.10.1515/cog-2012-0020

Cienki, A. 1998. Metaphoric gestures and some of their relations to verbal metaphoric expressions. In J. P. Koenig (ed.), *Discourse and cognition: Bridging the gap,* 189–204. Stanford, CA: CSLI Publications.

de Ruiter, J. P., Bangerter, A., and P. Dings. 2012. The interplay between gesture and speech in the production of referring expressions: Investigating the tradeoff hypothesis. *Topics in Cognitive Science* 4(2): 232–248. https://doi.org.10.1111/j.1756-8765.2012.01183.x

Enfield, N. J., Kita, S., and J. P. de Ruiter. 2007. Primary and secondly pragmatic functions of pointing gestures. *Journal of Pragmatics* 39: 1722–1741. https://doi.org/10.1016/j.pragma.2007.03.001

Goldin-Meadow, S. 2000. Beyond words: The importance of gesture to researchers and learners. *Child Development* 71(1): 231–239. http:doi.org/10.1111/1467-8624.00138

Gregersen, T., Olivares-Cuhat, G., and J. Storm. 2009. An examination of L1 and L2 gesture use: What role does proficiency play? *The Modern Language Journal* 93(2): 195–208. https://doi.org.10.1111/j.1540-4781.2009.00856.x

Gullberg, M. 1998. *Gesture as a communication strategy in second language discourse: A study of learners of French and Swedish.* Lund: Lund University Press.

Hanamoto, H. 2020. How do co-occurring speech and gestures express meaning in a dyadic face-to-face ELF interaction? In R. A. Giri, A. Sharma, and J. D'Angelo (eds.), *Functional variations in English: Theoretical considerations and practical challenges*, 61–82. UK: Springer. https://doi.org/10.1007/978-3-030-52225-4_5

Hanamoto, H. 2021. Improving spatial and temporal attention through the use of spontaneous gestures in dyadic English as lingua franca interactions. In C. Lee (ed.), *Second language pragmatics and English language education in East Asia*, 164–193. UK: Routledge.

Hostetter, A. B. 2011. When do gestures communicate? A met-analysis. *Psychological Bulletin* 137(2): 297–315. https://doi.org/10.1037/a0022128

Jefferson, G. 1984. On the organization of laughter in talk about troubles. In J. M. Atkinson and J. Heritage (eds.), *Structures of social action: Studies in conversation analysis,* 346–369. Cambridge: Cambridge University Press. https://doi/org/10.1017/CBO9780511665868.021

Kendon, A. 1980. Gesticulation and speech: Two aspects of the process of utterance. In M. R. Key (ed.), *The relationship of verbal and nonverbal communication and language*, 207–227. The Hague: Mouton De Gruyter Publishers. https://doi.org/10.1515/9783110813098.207

Kendon, A. 1990. Spatial organization in social encounters: The F-formation system. In A. Kendon (ed.), *Conducting interaction: Patterns of behavior in focused encounters*, 209–237. Cambridge: Cambridge University Press.

Kendon, A. 1994. Do gestures communicate? A review. *Research on Language and Social Interaction* 27(3): 175–200. http://doi.org/10.1207/s15327973rlsi2703_2

Kendon, A. 2004. *Gesture: Visible action as utterance.* New York: Cambridge University Press. http://doi.org/10.1017/CBO9780511807572

Kita, S. 2000. How representational gestures help speaking. In D. McNeill (ed.), *Language and gestur*e, 162–185. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511620850.011

Kita, S., and A. Özyürek. 2003. What does cross-linguistic variation in semantic coordination of speech and gesture reveal?: Evidence for an interface representation of spatial thinking and speaking. *Journal of Memory and Language* 48(1): 16–32. https://doi.org/10.1016/S0749-596X(02)00505-3

Kita, S., Özyürek, A., Allen, S., Brown, A., Furman, R., and T. Ishizuka. 2007. Relations between syntactic encoding and co-speech gestures: Implications for a model of speech and gesture production. *Language and Cognitive Processes* 22(8): 1212–1236. https://doi.org/10.1080/01690960701461426

Krauss, R. M., and U. Hadar. 1999. The role of speech-related arm/hand gestures in word retrieval. In L. S. Messing, and R. Campbell (eds.), *Gesture, speech, and sign*, 93–116. New York: Oxford University Press. https://doi.org/10.1093/acprof:oso/97801985524519.003.0006

Krauss, R. M., Chen, Y., and R. F. Gottesman. 2000. Lexical gestures and lexical access: A process model. In D. McNeill (ed.), *Language and gesture*, 261–283. Cambridge, UK: Cambridge University Press.

Lakoff, G., and M. Johnson. 1980. *Metaphors we live by*. Chicago, IL: Chicago University Press.

McNeill, D. 1992. *Hand and mind: What gestures reveal about thought.* Chicago: University of Chicago Press.

McNeill, D. 2005. *Gesture and thought.* Chicago: University of Chicago Press. https://doi.org/10.7208/chicago/9780226514642.001.0001

Meierkord, C. 2000. Interpreting successful lingua-franca interaction. An analysis of non-native/non-native small talk conversation in English. *Linguistik Online* 5: 1/00. https://doi.org/10.13092/lo.5.1013

Morrel-Samuels, P., and R. M. Krauss. 1992. Word familiarity predicts temporal asynchrony of hand gestures and speech. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 18(3): 615–622. https://doi.org/10.1037/0278-7393.18.3.615

Schegloff, E. A. 1992. Repair after next turn: The last structurally provided defense of intersubjectivity in conversation. *American Journal of Sociology* 97(5): 1295–1345. https://doi.org/10.1086/229903

Sime, D. 2006. What do learners make of teachers' gestures in the language classroom? *International Review of Applied Linguistics in Language Teaching* 44(2): 211– 230. https://doi.org/10.1515/IRAL2006.009

So, W. C., Kita, S., and S. Goldin-Meadow. 2013. When do speakers use gestures to specify who does what to whom? The role of language proficiency and type of gestures in narratives. *Journal of Psycholinguistic Research* 42(6): 581–594. https//doi.org/10.1007/s10936-012-9230-6

Torralbo, A., Santiago, J., and J. Lupiáñez. 2006. Flexible conceptual projection of time onto spatial frames of reference. *Cognitive Science* 30(4): 745–757. https://doi.org/10.1207/s15516709cog0000_67

Tuite, K. 1993. The production of gesture. *Semiotica* 93(1-2): 83–106. https://doi.org/10.1515/semi.1993.93.1-2.83

Wagner, P., Malisz, Z., and S. Kopp. 2014. Gesture and speech in interaction: An overview. *Speech Communication* 57: 209–232. https://doi.org/10.1016/j.specom.2013.09.008

Acknowledgments

The current research and the preparation of this paper was supported by the Japanese Government Grant-in-Aid for Scientific Research (No. 20K13113). The author thanks the students who made this research possible. Further acknowledgments are due to the audience for their very constructive comments on the author's oral presentation at LingBaW 2021, Poland (online) on October 15, 2021, and to the anonymous reviewers for their valuable feedback on earlier versions of this manuscript.

Appendix. Transcription conventions (Adapted from Jefferson 1984 and McNeill 2005)

:	extended sound or syllable
{	overlapping utterances with non-verbal actions
=	latched utterances
((	non-verbal actions
CHANGE	speech much louder than surrounding talk
(.)	short pause less than 1 second
[	overlapping utterances
→	feature of interest to analyze