The Achievement of Intersubjectivity through Embodied Completions: A Study of Interactions Between First and Second Language Speakers

1JUNKO MORI and 2MAKOTO HAYASHI

1University of Wisconsin-Madison, 2University of Illinois at Urbana-Champaign

This study examines casual interactions between first language (L1) and second language (L2) speakers of Japanese, paying special attention to the coordination of vocal and non-vocal resources that are brought to bear on the achievement of intersubjectivity. More specifically, this study investigates a practice of ‘embodied completion’ (Olsher 2004), namely the practice of deploying a partial turn of talk that offers a projectable trajectory of ongoing action and completing that action with a gesture or other embodied display. The participants’ conduct that precedes this embodied completion reveals the local processes used to evaluate, discover, and establish shared linguistic and non-linguistic resources in pursuing intersubjectivity. Further, the sequence of actions that follows the embodied completion provides an incidental, interactionally motivated opportunity for the L1 speaker to reformulate what the L2 speaker has said with a more sophisticated linguistic expression. Through the close analysis of two focal cases of embodied completion, which underscores the conversation analytic (CA) perspective of ‘competences as resources,’ this study explores the kinds of contributions that can be made by using CA’s explication of interactional details towards the understanding of language learning as it occurs in socially situated practices.

INTRODUCTION

In recent years, a number of studies conducted across disciplines, including communications, linguistics, psychology, and sociology, have demonstrated how different types of gestures are employed in talk-in-interaction to supplement or amplify vocal conduct and how these gestures are coordinated with, and situated in, the development of talk in talk-in-interaction (e.g. Goodwin 1981, 1995, 2000; Schegloff 1984, 1998; Heath 1986; Kendon 1990; McNeill 1992, 2000; Streeck 1993, 1994; Ford et al. 1996; Lerner 2002; Hayashi et al. 2002; Kita 2003; Hayashi 2003, 2005; among others). The increased attention to gestures and other non-vocal conduct has manifested itself as a promising future direction for studies in the field of second language acquisition (SLA) as well. As early as the 1970s, researchers such as Tarone (1977) and Færch and Kasper (1983) noted that mime and
gesture were used by second language (L2) learners as some of the most basic communication strategies to compensate for their lack of L2 proficiency, while Henzl (1979) and Hatch (1983) referred to the frequent use of gestures as a feature of talk addressed to L2 learners, so-called ‘foreigner talk’ (Ferguson 1971). However, it is during the last decade that gesture and other non-vocal conduct have become central objects of study in SLA (e.g. Adams 1998; Gullberg 1998; McCafferty 1998, 2002, 2004; McCafferty and Ahmed 2000; Lazaraton 2004; Negueruela et al. 2004; Olsher 2004; Belhiah 2005; with a small number of exceptions such as Neu 1990). The present study situates itself in this growing body of research that investigates the significance of gesture in the context of L2 use and its acquisition.

To be more precise, the current study sheds light on the contingent processes in which first language (L1) and L2 speakers coordinate talk and embodied action for the achievement of intersubjectivity, focusing on a practice of ‘embodied completion,’ that is ‘launching a turn at talk, and then at a point where some trajectory of the turn is projectable, ceasing to talk and completing the action that had been initiated by the particular turn through gesture or embodied display’ (Olsher 2004: 221). The close analysis of two cases of embodied completion, extracted from casual interactions among L1 and L2 Japanese speakers, demonstrates how the participants design these hybrid moves consisting of talk and gesture or embodied display according to their on-going estimation of shared resources. Further, the analysis of sequential developments that precede and follow the occurrences of embodied completion reveals how this practice not only facilitates comprehension but also triggers an incidental opportunity for the L2 speaker to learn more advanced linguistic forms. The remainder of this section provides a brief overview of recent developments in SLA studies on gesture, particularly focusing on those that emphasize the social aspects of the use of gesture. This overview clarifies the methodological and theoretical orientations that have shaped our study.

To our knowledge, the first comprehensive study of gesture as a communication strategy was conducted by Gullberg (1998), who investigated different types of gestures produced by learners of French and Swedish during a task of retelling a story that had been presented as a cartoon. The results of her study, generated by both quantitative and qualitative methods, indicated that the subjects produced more gestures when they were constructing narratives in their L2 than when they were performing the same task in their L1. Further, she reported that the gestures produced in their L2 narratives were not limited to mimetic gestures replacing speech and thereby solving lexical problems; they included more abstract types that are produced with speech for the purpose of handling discourse-related difficulties and metalinguistic commentary. Her study also explored how L1 speakers assess L2 learners’ gestures and how that assessment may affect their evaluation of L2 learners’ overall proficiency.
At around the same time, Adams (1998) conducted a study on gesture associated with ‘foreigner talk,’ by adopting a similar story-retelling task as the method for data collection. He examined how L1 speakers of English used gestures differently when addressing an L2 speaker of English (who was an L1 speaker of Korean) as opposed to addressing another L1 speaker of English. His quantitative study provided partial support for the hypothesis that native speakers would appeal to gestures more frequently in their interactions with nonnative speakers to make their messages more understandable. He also suggested the need to consider further if and how different levels of L2-speaking listeners’ proficiency may affect the ways in which L1-speaking narrators produce gestures.

These studies provided more vivid pictures of gestures produced by, or addressed to, L2 speakers than what had been speculated by their predecessors. The task design adopted by Gullberg (1998), Adams (1998), and several other recent SLA studies on gesture (following McNeil (1992) and his associates) offers the advantage of controlling the development of narratives and thereby making it easier to identify the connection between gestures and their intended meanings as well as to create taxonomies of the types and functions of gestures observed. However, such a design presents some limitations as well. One of the limitations, for instance, is their relatively weak account of interactional dynamics. Due to the task design which pre-assigns the division of labor, namely narrating and listening, the researchers’ attention tends to be unfairly directed towards the performance of narrators. Although Gullberg (1998) did analyze gestures used by L1-speaking listeners, the analysis constitutes an independent chapter rather than highlighting the moment-by-moment development of the interactional exchange that took place between the two parties.

As noted by Yule and Tarone (1991), and more recently by Firth and Wagner (1997), among others, although studies on communication strategies used by L2 speakers and foreigner talk performed by L1 speakers tend to have been developed separately, constituting seemingly different lines of research, the reality is that the features described in these studies do occur together, while shaping and reshaping one another, during a single interaction. Thus, rather than separately examining the two parties, that is narrators and listeners, or L1 speakers and L2 speakers, we should pose the question as to how the interlocutors build a common frame of reference which enables them to accomplish understanding when they discover they do not share common ground. A study of gestures, we believe, should also be conducted under this principle of investigating the co-constructional aspects of interaction.

One of the few SLA studies on gestures that takes such an intricate social dimension into consideration is McCafferty’s (2002) longitudinal study of interactions between a Taiwanese learner of English and an ESL instructor. McCafferty examines how these interactants, coming from different cultural and linguistic backgrounds, ‘had to co-construct their
own tool kit and practices, which included the use of gesture to reference artifacts and other features in the environment’ (McCafferty 2002: 200). In other words, the study investigates how the shared history of gestured signs established over time between the two, who share ‘no cultural tool kit,’ operates as a resource for enhancing comprehension and creating the zones of proximal development (ZPD) for language learning and teaching. Our study also pursues this line of inquiry: that is, it examines the processes of the interactants’ search for shared resources and the pivotal role served by gesture in these processes.

What distinguishes our study from McCafferty’s, however, is the theoretical and methodological orientation. McCafferty’s study is deeply rooted in the Vygotskyan tradition and motivated by the conceptual framework of the ZPD as presented by Newman and Holtzman (1993). Our study, on the other hand, was developed within the methodological framework of conversation analysis (CA) (cf. Sacks et al. 1974) and is concerned with the explication of publicly observable phenomena with reference to the fundamental interactional patterns (e.g. turn-taking mechanisms, sequential organization, preference organization, repair organization, etc.). Further, whereas McCafferty examined the diachronic establishment of shared resources between the Taiwanese learner and the ESL instructor, our study provides a synchronic, but detailed, analysis of the ongoing estimation and re-estimation of shared resources performed by L1 and L2 Japanese speakers, describing in detail how the participants themselves make sense of each other. The following quote from Heritage (1984) summarizes a primary aim of CA that the current study adopts:

Conversation analysis—like the other research streams of ethnomethodology—is concerned with the analysis of the competences which underline ordinary social activities. Specifically it is directed at describing and explicating the competences which ordinary speakers use and rely on when they engage in intelligible, conversational interaction. At its most basic, the objective is to describe the procedures and expectations in terms of which speakers produce their own behaviour and interpret the behaviour of others. (Heritage 1984: 241)

Of course, this premise concerning competences held among ‘ordinary speakers’ may need to be somewhat modified in the case of talk-in-interaction involving L2 speakers. While acknowledging the possible need for modification, our study essentially maintains this ethnomethodological principle in approaching the current data. Following the data analysis, we will revisit this issue as to how the understanding of competences may be applied to the analysis of L2 interaction and learning.

As mentioned earlier, the present study pays special attention to the practice of ‘embodied completion’ analyzed extensively by Olsher (2004). This phenomenon itself is not particularly new to researchers of L2 talk.
For instance, Tarone (1977: 197) introduced an example in which an L2 speaker said, ‘and everybody say,’ but, rather than verbally completing the sentence, clapped hands in order to express the idea that everybody applauds. Likewise, the aforementioned studies by Gullberg (1998) and McCafferty (2002) also refer to cases in which L2 speakers or the L1 speaker interacting with the L2 speaker leave sentences verbally unfinished and supply gestures instead. Olsher’s (2004) unique contribution, however, is his detailed CA account of this phenomenon that explores the mutual elaboration of talk and embodied practices in carrying out sequentially organized social actions. In his analysis of peer interactive talk in an English as a Foreign Language classroom in Japan, Olsher describes the emergence of these hybrid interactional moves with reference to their projectability of turns-in-progress as an essential resource, which allows for the understanding of embodied completions, as well as to the interactional contexts fostering these hybrid moves to be mobilized. Olsher also notes that while it might be tempting to claim that this practice is a coping mechanism uniquely employed by those who lack linguistic resources and fluency, occurrences of embodied completion are not limited to such seemingly deficit communicators. Introducing an excerpt of fully competent L1 speaker talk, Olsher argues that embodied completion is actually a generally available discursive practice for accomplishing various social actions, such as inviting the recipient of a criticism to actually voice the criticism or avoiding other delicate interactional matter. His classroom data indeed include a case in which an L2 speaker uses this practice, not necessarily due to her lack of L2 proficiency, but as a way of expressing her hesitance in complying with a request made by her peer.

Our study provides yet another context for the use of embodied completion. In our Japanese data, it is L1 speakers who utilize this practice when addressing L2 recipients. In these cases found in our data, the discursive practice of embodied completion, apparently available for L1 or L2 speakers of English, as well as Japanese, appears to be motivated by recipient-design considerations, that is the sensitivity and orientation to some specific features of the co-participants (cf. Sacks et al. 1974; Schegloff 1989). Namely, when facing co-participants who come from different cultural backgrounds and who do not appear able to manage the full range of L2 linguistic resources, participants need to evaluate constantly what kinds of vocal or non-vocal resources are shared or not shared among them as they proceed with the talk-in-interaction. Thus, the design of each turn can be viewed as the speaker’s demonstration of his or her understanding of the co-participants’ linguistic and other competences, which has been established through the course of talk-in-interaction leading up to the particular turn in question. By closely examining the moment-by-moment development of talk-in-interaction that precedes and follows the deployment of embodied completion, we will discuss how this practice not only facilitates understanding, but also provides an incidental, interactionally motivated
opportunity for L1 speakers to reformulate what the L2 speakers have said with a more sophisticated linguistic expression. The juxtaposition of the two versions of the verbalization of the embodied completion provides an opportunity for the L2 speaker’s noticing and potential learning of a refined linguistic form.

The following section provides a close analysis of two cases of embodied completion. The subsequent discussion addresses what kinds of implications can be drawn from the CA analysis of these cases with regard to perspectives on competences and language learning.

DATA ANALYSIS

Following a brief description of the participants and the setting of the interactions to be examined, this section first demonstrates that the deployment of embodied completion is an outcome of the L1-speaking participant’s on-going estimation of shared communicative resources with his L2-speaking co-participant who has, through the course of conversation, displayed a noticeable lack of linguistic competence in Japanese. In the second part of this section, we examine another similar case of embodied completion, with our focus this time on the participants’ conduct after the L1 speaker’s embodied completion. We argue that the sequence of actions following embodied completion can provide an incidental, yet interactionally motivated opportunity for presenting the L1 speaker’s reformulation of the L2 speaker’s verbal rendition of the embodied action.

Participants and setting

The two excerpts of interactions to be examined in this section were recorded at a student-organized weekly conversation table held at a local coffee shop in a university town in the Midwestern USA. The purpose of this event was to provide learners of Japanese with opportunities to practice using Japanese by engaging in casual conversations with native speakers of Japanese. There were no formal instructions as to how the participants should form groups or what they were supposed to talk about. Rather, they selected their partners on their own and talked about whatever topics they considered to be relevant or interesting at the moment. After obtaining the participants’ informed consent, a research assistant recorded their interactions with a handheld camera and a small wireless microphone that was placed on the table. The two L2 speakers, Alan and David, who appear in the following excerpts, had completed at least five semesters of college-level language instruction (approximately 275 contact hours) at a nearby university and were taking a sixth semester course at the time of recording. Neither one of them had spent an extended period of time living or studying in Japan prior to the dates of these recordings. The three L1 speakers in the first excerpt, Toru, Koji, and Yoshi, were students at an engineering university in Japan.
who were participating in a short-term (three- to four-week) English as a Second Language program offered at the university where Alan and David were studying Japanese. Like most Japanese students, these three had studied English since they were 12 or 13, but their oral communication skills were very limited judging from the ways in which they spoke English during these videotaped interactions. This was their first time participating in the conversation table. The two L1 speakers in the second excerpt, Taro and Sachi, on the other hand, were pursuing graduate degrees at the university and were regulars in this weekly gathering. Sachi was majoring in foreign language education, and Taro, who is the primary speaker in the segment analyzed below, was majoring in nursing. All the names of participants and towns are pseudonyms.

**Embodied completion as a recipient-designed practice**

In this subsection, we show how the deployment of embodied completion by an L1 speaker can be analyzed as the result of his recipient-design considerations. To support our argument, we closely examine the interactional processes through which the embodied completion emerges as a complex response to the local context of the immediately preceding sequence where the varying levels of linguistic knowledge among the L2-speaking co-participants become interactionally salient. Before probing the details of this preceding sequence leading up to the L1 speaker’s embodied completion, however, we first introduce the segment where the embodied completion in question takes place, so as to provide the reader with an initial understanding of how a gesture is deployed by the L1 speaker as a way to complete a social action initiated through the utterance-in-progress. In the following excerpt where Toru, an L1 speaker of Japanese, talks to two L2-speaking co-participants, Alan and David, consider how Toru begins his turn at talk with verbal expression, and then at a point where its grammatical structure is still incomplete, ceases to talk and completes the unfolding social action through the deployment of a hand gesture.

**Excerpt 1**

1 Toru: *nanode:: (1.0) hurui:: (.) kuruma wa:: (0.2)*
   so old car TP

2 *nedan ga:: [(0.4)*
   price SP [
   → [((moves open right hand downwards repetitively,
   → [ while gazing intently at Alan; Fig. 1))

   “So::, (1.0) as for o::ld (. ) ca::rs, (0.2)
   their pri::ces ((hand gesture follows))”

3 Alan: *u::::n* ((while nodding several times; David also nods))
   “I:: see::::”
4 David: *motto yasu[ī. (while nodding several times))
more cheap
“are cheaper.”

5 Toru: [yasui n desu. ((with an emphatic nod at yasui,
cheap N CP along with a single downward
“are cheap.” movement of an open right hand))]

In lines 1–2, Toru produces his talk slowly and carefully by stretching vowels
and pausing at the end of words and phrases, while gazing intently at one of
the two L2-speaking co-participants, Alan. The structure of the talk that he
presents in these lines is projectable as the topic-comment structure of the
following form:

$$[[\text{Topic NP}] - wa \ [\text{Subject NP}] - ga + [\text{Predicate}]]$$

where the topic NP (*hurui kuruma* ‘old cars’) is marked by the postpositional
topic particle *wa* and the subject NP (*nedan* ‘prices’) is marked by the
postpositional subject particle *ga*. This fairly common sentence structure
strongly projects that, after the subject NP is produced, there will be some
sort of predicate produced as the final element in Toru's utterance in
progress. It is in this projected slot for a predicate that Toru deploys a gesture
by making small repetitive downward movements with an open right hand
(see Figure 1).

That this gesture is designed by Toru as bringing to completion the action
that had been initiated by his partial turn seems to be indicated, at least in
part, by his intent gaze at Alan at the end of line 2. Such an intent gaze can
be regarded as a cue that responses by the recipients are relevant (as opposed
to, for example, gazing away from the recipients, a practice commonly
observed when the speaker displays his/her engagement in a ‘solitary’ word

![Figure 1: Toru’s embodied completion](image-url)
search in the midst of a not-yet-complete turn (Goodwin and Goodwin 1986; Hayashi 2003; Olsher 2004)). Indeed, Alan responds with an acknowledgment token (I see:::; line 3) and several nods, claiming his understanding of the meaning conveyed, whereas David, the other L2-speaking co-participant, responds with a linguistic expression (motto yasui ‘are cheaper’; line 4) that is designed to demonstrate (rather than merely claim) his understanding of the locally relevant meaning of the embodied display presented by Toru.

Now, how exactly does this embodied completion emerge as a locally appropriate form of conduct at the particular moment of the interaction in which it is produced? To answer this question, it is important to investigate the larger interactional context within which the turn in question is embedded. In what follows, we provide a close analysis of the following aspects of the context as potentially relevant to the participants’ conduct under investigation: (1) the broader activity context—a kind of telling—in which Toru’s contribution is produced and interpreted as a specific type of interactional move relevant to the unfolding course of the activity-in-progress; (2) the local context of the immediately prior sequence in which difficulties in achieving mutual understanding through linguistic communication have arisen as a practical problem that the participants needed to resolve in situ; and (3) the micro-level trajectory of the unfolding turn in question in terms of both vocal and non-vocal conduct produced by the speaker as well as the recipients. We argue that all of these aspects of the interactional context inform Toru’s on-going estimation of communicative resources shared with his co-participants and that the deployment of embodied completion is a situated outcome of such a moment-by-moment interactional calibration.

The broader activity context in which the segment in (1) is situated is that of a ‘telling’ by Toru (with assistance from the two other L1-speaking co-participants, Koji and Yoshi), addressed to Alan and David, about the prices of used cars in Japan as compared to those in the USA. This telling was initially prompted by David, who drew the attention of the L1-speaking co-participants to some advertisements of used cars in a local newspaper that he had in front of him, and asked them what they thought about the prices of used cars in the USA. In answering this question, Toru mentioned Japan’s sha-ken system, that is the government-enforced system that requires all automobiles to go through professional inspections at specified regular intervals. The aspect of the sha-ken system that is the most relevant to Toru’s telling delivered as an answer to David’s question is that the older a car gets, the shorter the interval becomes, and therefore that if one owns an old car, one ends up spending a large amount of money on this mandatory inspection. Thus, car owners in Japan, Toru states, tend to avoid keeping their cars more than 5–7 years and replace them with brand-new ones. This tendency results in a considerable lowering of the prices of older cars in Japan.
Toru’s explanation of this rather complicated situation becomes fairly lengthy, and during its course, the two L2 speakers display different levels of comprehension. While David keeps nodding during Toru’s talk and thereby indicates his understanding of the verbal description of the situation, Alan visibly shows his struggle in comprehending the explanation by producing a series of lateral head shakes. Alan’s struggle becomes even more pronounced in the following segment (which takes place immediately prior to the segment in (1) above), in which Alan explicitly says, *mada wakarimasen* (‘I still don’t understand’; line 11) when Toru asks him if he understood Toru’s verbal account.

Excerpt 2

1 Toru: ... *daitai::* (1.0) *gonen kara:* (1.0) *nananen de;*, usually 5.years from 7.years in
     “Usually::*, (1.0) in five to: (1.0) seven years:;”
2 David: *hai.* ((nods))
     “Yes.”
3 Toru: *kuruma o:* (. ) *kaeteshimaimasu.*
     car O change
     “they change their cars.”
4 (1.0)
5 David: *u::::::n.* ((nods several times))
     “Uh huh”
6 Koji: *muzukashikunai?* ((to Toru))
     difficult:NEG
     “Isn’t it difficult?”
7 Toru: *chotto muzukashii.*
     little difficult
     “a little difficult”
8 Toru: *waka- wakarimasu?* ((to Alan and David))
     understand
     “Do y- do you understand?”
9 10 David: *hai hai hai hai hai* ((nods several times))
     “Yes, yes, yes, yes, yes.”
11 → Alan: *a:::::: mada wakarimasen.*
     uh still understand:NEG
     “Uh:::::: I still don’t understand.”
12 Koji: *muzukashii [desu].*
     difficult CP
     “It’s difficult.”
In this segment, we can observe the struggle not only on the part of the L2-speaking recipients who are trying to understand the telling, but also on the part of the L1-speaking tellers who are trying to make their telling understood. For instance, in line 7, Koji explicitly assesses Toru’s design of talk as being too difficult. This assessment prompts Toru’s self-evaluation (line 8) and leads him to ask the recipients explicitly whether they have understood his explanation (line 9). Alan’s response to this question with a clear statement of non-understanding (line 11) is received with minimal acknowledgments by the L1-speaking tellers (lines 12 and 13), and the conversation falls into a momentary lapse (line 14), during which Toru assumes a ‘thinking posture’ (somewhat like the posture of The Thinker, a well-known sculpture by Rodin), which suggests that he has no immediate solution to the communicational difficulty at hand. When this ‘interactional dead air’ has started to prevail, David, the L2-speaking recipient who has been claiming understanding of Toru’s explanation, comes in and saves the moment. He offers a translation of the gist of the explanation into English for Alan (lines 15 and 17–18), to which Alan finally claims understanding (line 20). Thus, in this understanding–check sequence that takes place after Toru’s lengthy account of the sha-ken system and its consequences, the difficulties facing participants on both sides, that is Alan’s struggle in understanding what is being said and Toru’s struggle in making his telling understood, become interactionally salient and pose a practical problem that the participants need to resolve in situ as the interaction unfolds.

It is under the contingencies of such an interactional context that Toru’s production of a ‘hybrid construction’, that is a partial turn + embodied completion, takes place. We now examine the interactional processes through which this construction emerges as an outcome of an on-going,
finely tuned calibration of the communicative resources shared (or not shared) with these particular co-participants at this particular moment in interaction. Let us first describe how the participants’ bodily conduct is coordinated with that of the other’s in the evolving production of the partial turn by Toru (nanode:: (1.0) hurui:: () kuruma wa:: (0.2) nedan ga:: ‘So::, (1.0) as for o::ld (. ) ca::rs, (0.2) their pri::ces’).

Excerpt 3

At ‘A’ above

As Toru produces the discourse connective nanode (roughly equivalent to so in English, used turn-initially to present an upshot of what the speaker has been saying), he starts to extend his right arm and points his index finger (and directs his gaze) at the advertisements of used cars in the local newspaper that David has in front of him. Note that these advertisements are the ones to which David had drawn everyone’s attention when he asked his original question about used-car prices that led Toru to initiate this whole ‘telling’ activity. Thus, with the concurrent deployment of these bits of vocal and visual conduct turn-initially, Toru displays to his co-participants that the interactional move he is about to launch in the unfolding turn involves (1) a return to the topic of used-car prices after the lengthy and complicated explanation of the sha-ken system, and (2) a presentation of an upshot of the preceding explanation, that is the consequence of the existence of the sha-ken system to the prices of used cars in Japan.¹

Just as Toru’s index finger reaches the newspaper in front of David at de:: in nanode::, Alan and David start to shift their gaze from Toru’s face to where his index finger is placed on the table, and the three participants enter into a framework of mutual orientation with their common focus of attention on the advertisements being pointed at. This suggests that Toru’s pointing is
indeed seen and attended to by his co-participants as a relevant part of the communicative conduct in the unfolding course of the social action-in-progress.

At ‘B’ above

As Toru moves on to produce hurui:: (‘old’), he looks up and brings his gaze to Alan, the co-participant who displayed considerable struggle in understanding Toru’s verbal explanation in the prior sequence. Toru maintains his intent gaze at Alan as he goes on to produce kuruma wa:: (‘ca::rs’). This sustained gaze suggests that he is building his turn as specifically addressed to Alan, perhaps to make sure that Alan understands his utterance-in-progress this time.

After Toru brings his gaze to Alan during hurui::, Alan also starts to look up and shifts his gaze from the used-car advertisements on the table to Toru’s face. Thus, at the micro-pause after hurui::, Alan enters into a state of mutual gaze with Toru, which continues throughout the remainder of the unfolding turn. Alan not only displays his attention to Toru’s talk through his gaze, as Toru produces wa:: in kuruma wa::, Alan deploys a series of slight head nods, and thereby indicates his involvement in (and perhaps claims his understanding of) Toru’s utterance-so-far. Given that Toru appears to be pursuing Alan’s understanding of his utterance as stated above, this claim of involvement (and possibly understanding) from Alan may provide some basis for Toru to estimate that some level of understanding of his talk-so-far has been reached.

At ‘C’ above

During the 0.2-second pause after huruma wa::, Toru starts to slightly retract his extended right arm that had been laid on the table, while at the same time raising his open right hand (with the palm facing down) to the chest level. When the open hand reaches the chest level as he vocally produces dan in nedan (‘prices’), Toru momentarily ‘freezes’ the hand movement there until the end of the stretched articulation of the subject particle ga::: The sequence of these hand and arm movements can be seen as what McNeill (1992) calls the ‘preparation phase’ of gesticulation (the slight retraction of the arm and the raising of the hand) followed by the ‘pre-stroke hold’ (the freezing of the hand at the chest level). These preparatory movements strongly project a subsequent deployment of a gesture stroke, that is the peak of gesticulation. Toru does all this preparatory work while keeping his gaze on Alan, and Alan also maintains a reciprocal gaze throughout. Thus, the two participants find themselves in an opportune moment for gesticulation—a moment at which a gesture about to be produced is likely to be delivered in mutual gaze with its addressee.\(^2\)
On the linguistic side of Toru’s conduct, the progressive realization of grammatical structure in Toru’s on-going utterance narrows down possible trajectories of the remainder of the turn, and it makes the interpretation of what is going to be produced in the next moment highly projectable. That is, the structure of Toru’s utterance-so-far, that is a topic NP (hurui kuruma wa ‘as for old cars’) followed by a subject NP (nedan ga ‘their prices’), makes it strongly projectable that the next element in the turn will convey the meaning of ‘being reduced/lowered’.

So far, we have examined in detail what transpired both vocally and visually during Toru’s delivery of the partial turn prior to the deployment of embodied completion. Based on this examination, we can summarize and highlight potentially relevant aspects of the interactional contingencies Toru faces at the end of the partial turn which might influence his on-going calibration of the communicative resources which he shares (or does not share) with his addressee Alan:

- Alan can be assumed to possess the grammatical knowledge in Japanese that nedan ga (Noun + Subject Particle) will be followed by some kind of predicate. The topic-comment structure of the form [TOPIC NP-wa [SUBJECT NP-ga + PREDICATE]] is one of the most common and basic sentence patterns in Japanese, and, although Alan’s knowledge of Japanese may be limited, through the course of the conversation he has shown himself to possess proficiency in the language far beyond such a basic level.

- Given that (1) Alan achieved, thanks to David’s translation, a basic understanding of Toru’s prior telling about the sha-ken system and the related tendency for people to avoid buying older cars in Japan; that (2) Alan’s visual conduct during the production of the partial turn suggests that he is following Toru’s utterance-so-far; and that (3) Alan must have a basic understanding of how a market economy works, he can be assumed to be in a position to anticipate what kind of concept will be expressed by the projected predicate regarding the prices of old cars in Japan.

- Though Alan may have a basic idea about the kind of concept to be expressed by the projected predicate, it is not at all certain whether he knows specific lexical items in Japanese that would complete the utterance-so-far in a contextually appropriate manner, such as sagaru (‘go down’), yasuku naru (‘become cheaper’), among others.

- It seems safe to assume, however, that Alan shares a basic understanding of a simple descriptive gesture of a hand moving downwards repetitively as an expression of some sort of ‘gradual descent,’ and that he can tie this basic understanding to the locally relevant meaning of the gesture at that moment, that is the lowering of used-car prices.

It is perhaps based on these considerations of the estimated knowledge and abilities of the recipient that Toru resorts to non-linguistic modality in order to facilitate and achieve successful communication. Rather than risking yet
another non-understanding by using a linguistic item that may not be shared by Alan, Toru makes use of a very basic kind of metaphorical gesture to complete the interactional move he has been building in the turn-in-progress, that is a presentation of an upshot of the telling-so-far. Thus, the embodied completion observed here, we argue, is an outcome of recipient design motivated by the speaker’s on-going calibration of what communicative resources are shared (or not shared) with his particular addressee.

Before we close this subsection, let us briefly describe what happens as Toru deploys the embodied completion. As soon as he starts to move his hand, Alan starts to nod in such a way that the timing of his nods synchronizes somewhat with the ‘beats’ of the repetitive hand movements of Toru’s gesture. As he nods, then, Alan produces a response token that claims some level of understanding (u::n ‘I:: see::’), while starting to move his gaze away from Toru, thereby exiting from the state of mutual gaze with him. This gaze shift suggests Alan’s understanding that Toru’s execution of the act of presenting an upshot of his telling-so-far is now complete. Indeed, Toru winds down the gesture as Alan moves his gaze away, and synchronizes the cessation of the hand movement with the end of Alan’s response token, as if to indicate that the interactional move that he has been launching is indeed complete now. It thus appears that the multi-modal hybrid construction that Toru produced as a result of his situated calibration of shared communicative resources has successfully brought about intersubjective understanding among the participants.3

Third-turn receipt as an interactionally motivated place for reformulation of L2 talk

Let us examine another case of embodied completion, with a focus now on the participants’ conduct after the embodied completion. We show that the sequence of actions following embodied completion can provide an incidental, yet interactionally motivated opportunity for presenting a reformulation of L2 learners’ verbal rendition of embodied action.

The following excerpt is taken from a longer sequence in which the participants (Alan, Taro, and Sachi) discuss people’s attitudes towards underage drinking in certain towns in the Midwest. The segment starts with Taro’s response to Alan’s prior comment (not shown in the excerpt below) that people have a strict attitude against underage drinking in the town of Adams. Taro initiates his talk in a manner that prefigures a disagreement (e demo: ‘Oh but’; line 1) and brings this disagreement to completion with an embodied display (line 3). Consider how Alan responds to Taro’s embodied completion (lines 4 and 7) and how Taro responds to Alan’s response (lines 8–9).
Excerpt 4

[establishes mutual
gaze with Alan]]

1 Taro: e demo: footo dikuson [no baa (e) [itta toki::,
          oh but Fort Dixon LK bar (to) went when
         “Oh but when I went to a bar in Fort Dixon,’’
          ]] 

2 Alan: [sore-
         that
         ]

3 Taro: [Aidii::
         (0.5)
         ID
         “ID:::”
         ]

4 Alan: [uha hah
         “uha hah’’

5 Sachi: uha ha ha [ha
         “uha ha ha ha”
          ]

6 Taro: [zenzen
         at.all
         “at all”
          ]

7 Alan: [tsukawanai. ((brings gaze to Taro))
         use:NEG
         “(you) don’t use.”

8 Taro: tsukawanai. datte min- aidii:: mi- misero to
         use.NEG because ID show QT

9 Taro: iwarenai
tell:PASS:NEG
         “(You) don’t use. Because uh you are not told you need to show
your ID.”

In response to Alan’s reference to the strict attitude against underage drinking in Adams, Taro mentions the time when he went to a bar in Fort
Dixon (line 1), a town not far from Adams. After achieving mutual gaze with Alan at the end of line 1, Taro shifts his gaze to Sachi and produces the word *aidii* (‘ID’; line 3), while deploying a gesture with both hands that depicts the act of pulling out an ID card from a wallet. The trajectory of Taro’s utterance-so-far and the projected action in progress (i.e. disagreement) make it highly likely that some sort of predicate, a negative predicate, to be precise, will be produced after the word *aidii*. Instead of providing a linguistic expression, however, Taro deploys an embodied completion through lateral head shakes produced successively for several times, while bringing his gaze back to Alan and thereby showing that Alan’s response is relevant (line 3).

In response to this embodied display by Taro and his gaze, Alan first reciprocates head shakes accompanied by laughter, while gazing away from Taro (line 4). A short moment later, Alan brings his gaze back to Taro and produces a linguistic expression, *tsukawanai* (‘you don’t use’; line 7), to offer his understanding of the locally relevant meaning of Toru’s embodied completion in line 3. Alan’s proffer of a candidate understanding of what Taro’s embodied action conveys makes Taro’s response (acceptance/rejection) relevant in the next slot in the sequence. Note then what Taro does in the subsequent turn (lines 8–9). He first accepts Alan’s candidate understanding by repeating the same lexical item *tsukawanai*; however, he goes on to provide an alternative way to verbalize the meaning of his embodied display, *aidii misero to iwarenai* (‘you are not told that you need to show your ID’). Compared to the expression produced by Alan, this alternative linguistic expression contains more ‘advanced’ and therefore more ‘challenging’ elements of Japanese grammar, such as a verb form conjugated with a passive morpheme, a subordinate clause embedded in a matrix clause, etc., as seen in the following:

```
subordinate clause             matrix clause
[aidii  mise-ro ]-to  iw-are-nai
    ID  show-IMPERATIVE  COMPLEMENTIZER  say-PASSIVE-NEGATIVE
```

‘One is not told that s/he needs to show his/her ID.’

Thus, Alan is presented with different versions of verbalization of the same embodied display side by side, that is a simple predicate produced by himself and a more sophisticated one produced by Taro, as the sequence of actions moves forward in the interaction.

It is important to register here that the difference between the two versions of ‘speech affiliate’ of Taro’s embodied action concerns not only the level of morphosyntactic complexity, but also the level of specificity of meaning that is relevant to the local course of action. That is, while Alan’s version, *tsukawanai* (‘one doesn’t use (his/her ID)’) simply indicates a customer’s choice of not using an ID, Taro’s version, *misero to iwarenai* (‘one is not told that s/he needs to show (his/her ID)’), highlights the fact that bar employees
do not ask customers to show their IDs. Given that the point of Taro’s
disagreement with Alan’s prior comment concerns how relaxed bar
employees were about enforcing the law regulating underage drinking in
the region under discussion, the important issue is not the customers’ choice
regarding the use of IDs, but the lack of enforcing the law on the part of the
alcohol providers. Thus, the precision in meaning brought about by Taro’s
version can serve to provide a learning opportunity for Alan to link the
contextually relevant aspects of the meaning of the gesture, on the one hand,
with the exact linguistic expression that conveys them, on the other.

What the preceding discussion suggests is that, while embodied completion
can provide a useful communicative resource to accomplish some level of
intersubjective understanding when interacting with co-participants who
may not share full linguistic competence in a language (as discussed in the
preceding subsection), the level of specificity (and/or accuracy) in meaning
that can be communicated through embodied action may be limited.
The third-turn receipt then provides the L1 speaker with an opportunity to
mobilize linguistic resources to rework and refine the mutual understanding
once achieved through embodied action, and this can provide the
L2-speaking recipients with exposure to a more ‘advanced’ linguistic
expression that elaborates on the idea which was conveyed in a more
holistic manner through bodily conduct. Thus, here we observed a specific
moment in which participants’ engagement in real-life interaction can
occasion a critical opportunity that may facilitate language learning in situ.

The case discussed here may remind some readers of ‘recasts’ (Doughty
1994; Lyster and Ranta 1997; Lyster 1998; Mackey and Philp 1998; Ohta
2000, 2001; Nicholas et al. 2001). The definitions of recasts vary among
researchers, but they are generally understood as restatements of language
learners’ utterances, which provide expansion or implicit correction while
maintaining the semantic content. Previous SLA studies of recasts tend to
view recasts as a kind of ‘corrective feedback’ and investigated their
effectiveness for the learners’ interlanguage development. The present case,
on the other hand, shows how an action that may be seen as an example of
recast (but not necessarily as ‘corrective feedback’ aiming at an L2 learner’s
linguistic error) is produced in situ as the participants work towards achieving
intersubjectivity through embodied completion and beyond (cf. Jefferson
1987). In other words, what appears to be an example of recast is designed
here to serve the interactional project of re-framing the ratified talk.

SUMMARY AND DISCUSSION

This study explored the dynamic relationship between gesture and other
nonvocal conduct and talk, by paying special attention to the interactional
contingencies that triggered the coordination of these different modes of
communication. More specifically, it demonstrated how a practice of
embodied completion took place in interactions between L1 and L2 speaking
participants, who were attempting to achieve intersubjectivity by relying on what they had come to view as shared resources. While the L2 speakers exhibited their limitation in linguistic sophistication, they also demonstrated their orientation towards the unfolding sequential and syntactic organizations. Such a demonstration situated the L1 speakers’ use of gestures and gaze directions during the production of earlier components of utterances as well as their eventual deployment of embodied completion. By undertaking the embodied completions, the L1-speaking participants could avoid linguistic expressions which may not have been accessible for the L2-speaking co-participants. The L1-speaking participants’ use of gestures, which signify particular ideas in a ‘global/synthetic’ manner as opposed to linguistic expressions that are more ‘segmented/analytic’ (McNeill 2000), enabled the participants to accomplish an approximate understanding between themselves, which was demonstrated by the L2-speaking participants through their production of corresponding gestures as well as verbalizations at their level of limited linguistic competence. The accomplishment of the approximate understanding, then, seems to have provided an opportunity for the presentation of refined linguistic expressions that further specified the exact ideas originally expressed through the gestures. The juxtaposition of the gesture used for an embodied completion and the varying versions of its verbalization can be considered as an opportunity for the L2-speaking participants to examine the relationship between the linguistic forms and the ideas expressed by them in their situated social actions.

Obviously, the current study, which offers an analysis of just two cases of embodied completion, does not intend to make any generalizable claim on occurrences of this particular practice in interactions involving L2 speakers in their aggregate. Indeed, as Olsher (2004) points out, this practice can be deployed in various sequential units and for accomplishing various social actions. The two cases introduced here also demonstrate that the practice can be embedded in the action of conveying new information and confirming the recipient’s understanding, or in the action of delivering a disagreement. However, in both cases, we can recognize that the emergence of this practice seemed to indicate the speakers’ recipient design considerations, particularly their awareness that the recipients were not fully competent speakers of Japanese. These cases also demonstrate that L1 speakers’ third-turn receipts following the L2 speaker’s verbalization of gestures provide important opportunities for presenting different versions of linguistic expressions for the meanings indicated through the gestures.

Through the close observations of hybrid moves, our study emphasized the perspective of ‘competences as resources,’ which is a fundamental working principle in CA, rather than that of ‘competences as objectives of learning,’ which has been more frequently highlighted in SLA research. According to Schegloff (1989), in the process of first language acquisition, children ‘learn to deal with the moment-to-moment contingencies of life in interaction, and the details of language use and conduct, in the moment-to-moment
contingencies of life in interaction, with their deployments of language and other conduct’ (Schegloff 1989: 152, emphasis original). These competences serve as both resources for, and objects of, learning at the same time. It is our belief that this understanding should apply to the process of second language acquisition as well, although in the case of adult L2 learners, we can assume that, in spite of their lack of linguistic competence specific to the L2, these learners bring a full range of fundamental competences, gained through their first language acquisition process, to the site of the L2 interaction. Recent CA studies of languages other than English have demonstrated that the kinds of fundamental organization (e.g. turn-taking, sequence, repair, preference, etc.) described by CA studies appear to operate across different languages and cultures (e.g. Sorjonen 1996, 2001; Hakulinen 2001; on Finnish; Egbert 1996, 1997a, 1997b; Golato 2000, 2002; on German; Mori 1999; Tanaka 1999; Hayashi 2003; on Japanese; Kim 1999, 2001; Park 1999, 2002; on Korean; Wu 2004 on Mandarin; to name a few). And if so, it is not difficult to imagine that the participants of intercultural interactions will share at least some of these fundamental, seemingly universal, patterns of interaction, and rely on them to accomplish interactional moves. And, as they interact with each other, they can constantly evaluate, discover, and formulate what are shared and not shared resources. The development of L2 specific linguistic competence, we believe, can be enhanced through such a maximization of available resources shared among the interactional participants. The data examined in the current study do not necessarily exhibit the participants’ orientations towards the interaction as instructional, nor do they present evidence as to whether or not the L2 speakers had indeed acquired the new forms and expressions presented to them. Nevertheless, the data clearly show that the sequenced actions of L1 speakers’ hybrid moves, L2 speakers’ verbalization of gestures demonstrating their understanding, and L1 speaker’s ratification and elaboration of the L2 speakers’ verbalization in third-turn receipts constitute interactionally situated learning opportunities. And the embodied completion plays a pivotal role in furthering this process. It must also be noted that the processes explicated in this study indicate that learning is not uni-directional (i.e. L2 speakers trying to learn to communicate better in the L2), but is bi-directional (i.e. both L1 and L2 speakers trying to find ways to accomplish intersubjectivity).4

What is described in this article can be linked to several different theories that have been influential in the SLA field. Recipient design considerations and the formation of hybrid moves may be approached from the perspective of language socialization (cf. Schieffelin and Ochs 1986), or the creation of the ZPD, scaffolding, or mediation (cf. Vygotsky 1978; Newman and Holtzman 1993). As mentioned earlier, what we observed in the third-turn receipt can be also considered an example of recast and discussed from the Interactional Hypothesis perspective (cf. Doughty 1994; Lyster and Ranta 1997). However, it is not our concern, as CA researchers, to judge which of these theories would explain better than others the potential learning
exhibited in these cases (although we would be happy to see such a discussion or a potential collaboration, prompted by the current study). What we do hope to have shown, however, is an illustration of how much we, the researchers, can learn about L2 interaction and learning from an analyst’s engagement in the moment-by-moment re-enactment of the processes in which the L1 and L2 participants themselves analyze each other’s demonstrated competences and design their moves by employing vocal and non-vocal resources. CA’s contributions towards SLA, we think, exist as an analytical potential that can provide concrete, empirical evidence that complement and/or question theoretical, hypothetical discussions. Thus, the data-driven approaches and theory-driven approaches should continue to maintain a healthy tension between them to advance our understanding of L2 use and acquisition.

Final version received 1 June 2005

ACKNOWLEDGMENTS

Earlier versions of this article were presented at the International Conference on Conversation Analysis in Copenhagen Denmark (May, 2002), the SLATE Lecture Series at the University of Illinois at Urbana-Champaign (April, 2003) and the Sociocultural Theory and Conversation Analysis in Second Language Acquisition Research Forum at the University of Southern Denmark (October, 2003). The comments we received from the audience at these meetings were invaluable. We are also grateful for constructive feedback provided by Numa Markee, Richard F. Young, Tobias Barske, Emma Betz, Patricia Yu, Elizabeth Miller, the three anonymous reviewers, and the co-editors of the journal, Gabriele Kasper and Guy Cook, who carefully read earlier versions of our article.

APPENDIX: TRANSCRIPTION CONVENTIONS AND ABBREVIATIONS

1. Transcript symbols

[ ] The point where overlapping talk and/or gesture starts
(0.0) The point where overlapping talk and/or gesture ends
(.) length of silence in tenths of a second
Micro-pause less than 2/10 of a second
Underlining relatively high pitch
CAPS relatively high volume
:: lengthened syllable
- cut-off; self-interruption
= ‘latched’ utterances
?/.!, rising/falling/continuing intonation respectively
! animated tone, not necessarily an exclamation
( ) unintelligible stretch
(word) transcriber’s unsure hearings

JUNKO MORI and MAKOTO HAYASHI 215
transcriber’s descriptions of events, including nonvocal conduct

hh audible outbreath
.hh audible inbreath
(hh) laughter within a word
< increase in tempo, as in a rush-through
○○ a passage of talk quieter than the surrounding talk

2. Abbreviations used in the interlinear gloss

<table>
<thead>
<tr>
<th>CP</th>
<th>various forms of copula verb be</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>nominalizer</td>
</tr>
<tr>
<td>O</td>
<td>object particle</td>
</tr>
<tr>
<td>QT</td>
<td>quotative particle</td>
</tr>
<tr>
<td>TP</td>
<td>topic particle</td>
</tr>
<tr>
<td>LK</td>
<td>nominal linking particle</td>
</tr>
<tr>
<td>NEG</td>
<td>negative morpheme</td>
</tr>
<tr>
<td>PASS</td>
<td>passive morpheme</td>
</tr>
<tr>
<td>SP</td>
<td>subject particle</td>
</tr>
</tbody>
</table>

NOTES

1 See Goodwin (2003) for an illuminating account of how the ‘meaning’ of a pointing gesture is construed through a juxtaposition of multiple semiotic resources, including the accompanying talk, the surrounding material environment, and the sequential and activity contexts in which the gesture is situated.

2 In order for a gesture to count as an interactionally meaningful communicative behavior, it must be seen by the recipient. It is recurrently observed that the producer of a gesture engages in preparatory work to make sure that the recipient is visibly positioned to attend to a forthcoming gesture (Streeck 1993; Goodwin 2000).

3 As Excerpt (1) shows, the sequence described so far does not end there and is continued with David’s verbal rendition of Toru’s gesture, followed by Toru’s response to it. While these latter bits of conduct are significant aspects of the process of achieving intersubjective understanding (as discussed in Mori and Hayashi 2002, 2003), for reasons of space we cannot discuss them in any detail here. The significance of the ‘third-turn receipt’ by the producer of the gesture will be explored (with a different example) in the next subsection.

4 The practice observed in this study may remind some readers of ‘designedly incomplete utterances (DIUs)’ discussed by Koshik (2002). There are in fact certain similarities in terms of the incompleteness of verbal turns. However, DIUs are used to elicit knowledge displays in error correction sequences in instructional settings, clearly indicating their producers’ orientation towards their institutional roles. On the other hand, embodied completions discussed here are not associated with institutional roles or purposes. Rather they contingently emerge in the processes of searching for resources to facilitate mutual understanding.
REFERENCES


Mori, J. and M. Hayashi. 2002. ‘What can collaborative completion tell us about learning a foreign language?’ Paper given at the International Conference on Conversation Analysis (ICCA-02), Copenhagen, Denmark.


