Available online at www.sciencedirect.com

ScienceDirect

Journal of Pragmatics 70 (2014) 68–85

journal of PRAGMATICS

www.elsevier.com/locate/pragma

Mimicked gestures and the joint construction of meaning in conversation

Kawai Chui *

Department of English & Research Center for Mind, Brain and Learning, National Chengchi University, Taiwan, ROC

Received 30 October 2013; received in revised form 24 May 2014; accepted 13 June 2014

Abstract

The current study investigates a complete course of action for the joint construction of meaning and the way mimicked gestures are used along with speech to accomplish the joint action in Mandarin Chinese conversation. The domain of analysis is a stretch of talk that encompasses the beginning till the end of the joint action during which similar gestures are produced by different speakers across turns. Within the stretch of talk, the beginning of the joint action is the 'presentation phase' during which a speaker presents meaning. A variety of situations were found to prompt another participant to jointly create meaning. The end of the joint action is the 'completion phase', during which the new meaning is recognized and the collaboration ends. In between is the 'collaboration phase' during which the joint action starts and develops with the use of cross-modal resources. In conversation, one way to accomplish the joint action is by the use of gestural repetition with slight modification as in a discussion about size. For other types of semantic information, the involvement of speech and gesture is more frequent, in that the second speaker mimics the gesture of the previous speaker to form a semantic foundation shared by the participants; and the second speaker conveys new meaning with a new lexical expression on the basis of the semantic common ground. The use of cross-modal resources thus facilitates the simultaneous realization of shared knowledge in gesture and new meaning in speech within a clausal unit.
© 2014 Elsevier B.V. All rights reserved.

Keywords: Mimicked gestures; Joint action; Meaning; Conversation

1. Introduction

In daily conversation, a common use of language is to construct meaning for communication, which can be accomplished via the collaboration between participants. A joint action such as this ''is carried out by an ensemble of people acting in coordination with each other'' (Clark, 1996:3). In order to understand how people collaborate to establish meaning together, it is first necessary ''to develop a detailed account of a course of action to understand what part any utterance plays within it'' (Lerner, 2002:249). In the past studies of the grounding of meaning for mutual understanding (Clark and Wilkes-Gibbs, 1986; Clark and Schaefer, 1987; Clark and Brennan, 1991; Clark and Krych, 2004), two phases of the joint action were proposed, namely the 'presentation phase' during which a referent is uttered by a speaker, and the 'acceptance phase' during which the referent is acknowledged by the other speaker. While such pattern is not directly subject to speakers' collaboration to construct meaning together, a complete course for such particular type of joint action has not yet been well investigated. Thus, the current study aims to examine a complete course of the action for the joint construction of meaning, which is preliminary and crucial to the investigation of the occurrence of mimicked gestures.

* Correspondence to: Department of English, National Chengchi University, No. 64, Sec. 2, ZhiNan Rd., Wenshan District, Taipei City 11605, Taiwan, ROC. Tel.: +886 2 29393091x88088; fax: +886 2 2939 0510.
E-mail address: kawai@nccu.edu.tw.

http://dx.doi.org/10.1016/j.pragma.2014.06.005
0378-2166/© 2014 Elsevier B.V. All rights reserved.

In that joint actions are social interaction, not only language, but also ''the body-in-action is available as a situated social resource'' (Lerner, 2002:250). The present study focuses on hand gestures, not only because the use of hands and arms along with speech is prevalent in daily communication (McNeill, 1992, 2000; Goldin-Meadow, 1999; Kendon, 2004), but also because gestures play a role in the joint creation of meaning. While co-speech gestures are largely spontaneous and speakers usually have their own ways to depict meaning based on their personal experiences and conceptualization, there is also a type of gesture that repeats the manual configuration which has previously been produced by another speaker for the same referent. This kind of repetition is 'gestural mimicry' (Kimbara, 2006:41) and involves ''the recurrence of the same or similar gesture across speakers.'' Gestures of this kind are also called 'return gestures' (de Fornel, 1992), 'gestural rephrasings' (Tabensky, 2001), 'mimicking gestures' (Kimbara, 2006) or 'mimicked gestures' (Parrill and Kimbara, 2006; Holler and Wilkin, 2011). Here, the term 'mimicked gesture' is used.''

Different kinds of mimicry data have been elicited by use of different methodologies in the previous studies, including data from videophone conversations (de Fornel, 1992), discussions on assigned topics between participants who had not met each other before (Tabensky, 2001), cartoon narrations and casual Japanese conversation (Kimbara, 2006), joint descriptions of video stimuli (Parrill and Kimbara, 2006; Kimbara, 2008), and face-to-face dialogs in which participants talked about a set of geometrical figures as stimuli (Holler and Wilkin, 2011). Studies using these various kinds of data provided different results about the occurrence of mimicked gestures. Parrill and Kimbara's (2006:165) experiments found that observers were sensitive to mimicry, in that ''participants who observed mimicry in gesture produced more of the mimicked features.'' Kimbara (2008), by manipulation of the visibility of the speakers in the joint description of a cartoon, and Holler and Wilkin (2011), by doing the same in a referential communication task, consistently found that a greater number of mimicked gestures were produced when the participants could see each other. In other empirical research, the use of recurrent gestures demonstrated attentiveness, strong involvement in the interaction, understanding and acknowledgment, alignment or heckling, and the joint establishment of meaning (de Fornel, 1992; Goodwin and Goodwin, 1992; Tabensky, 2001; Kimbara, 2006; Holler and Wilkin, 2011).

In the studies of the joint creation of meaning, Tabensky (2001) presented two examples of gestural rephrasings and supported de Fornel's (1992) claim that the use of mimicked gestures displayed participants' strong involvement during speaking. Five examples were discussed in Kimbara (2006) to show that gestural repetition made salient the foregrounded aspects of meaning, reshaped the co-expressivity of the interlocutors' speech and gesture, and accomplished the co-construction of utterance units. Finally, quantitative evidence can be found in Holler and Wilkin (2011) to manifest that mimicked gestures can be used for 'presentation', 'acceptance', and 'displaying incremental understanding'. However, as mentioned above, the data in Tabensky (2001) were from conversations between participants who did not know each other before the recording; three topics were assigned from which participants chose one for discussion. The data in Kimbara (2006) were mainly elicited from stimuli for joint description, as the participants first watched cartoon clips and then described the content together. In Holler and Wilkin (2011), although the data were more natural as participants engaged in face-to-face dialogs, the interaction was still task-based, in that ''the content of their talk focused on referents to concrete concepts as well as spatial information, and the exchange was structured in the sense that those participating in it adopted particular roles'' (Holler and Wilkin, 2011:136). Thus, the remaining question is: Do mimicked gestures play similar roles in naturally-occurring daily conversation?

In the present study, the occurrence of mimicked gestures along with speech during the course of action for the joint construction of meaning is examined in the most fundamental type of talk-in-interaction – daily face-to-face conversation (Sacks et al., 1974; Clark, 1996; Stivers et al., 2009). In this casual and spontaneous type of interaction, participants are free to talk about any topics of interest in their own way and develop joint actions without assignment of the topic or the use of video stimuli. The sequential turns in conversational interactions are also pertinent to the understanding of the participants' collaboration in joint actions.

In brief, the current study aims to investigate a complete course of action for the joint construction of meaning and the occurrence of mimicked gestures, to understand the circumstances under which meaning is jointly constructed, and the way that linguistic and gestural resources are employed to achieve the joint action in conversational discourse. It will be shown that the use of mimicked gestures for the collaborative act in daily conversation differs from their occurrence in reference communication tasks as found in previous studies. The findings will contribute to the understanding of the use of mimicked gestures as they occur along with speech in conversational discourse and the role of gesture in communication.

The next section introduces the data for the present study. Section 3 is the analysis of the course of the joint construction of meaning and the occurrence of mimicked gestures in conversation. Section 4 is the general discussion and conclusion.

2. Data and methods

The data for this study consist of daily face-to-face casual conversations among adult native speakers of Mandarin. The participants were recruited to hold a conversation with their friends, family members or colleagues who knew each

other. Participants gave written consent for the data to be accessed online.[1] All of the participants were paid, and they were told that they were participating in research on conversation, but gestures were not mentioned. For each recording, the participants chose a place where they could talk in a leisurely manner, such as a classroom, students' lounge, dorm room, or living room. Participants were free to find and develop topics of common interest, and were filmed for approximately an hour with a visible camera. One stretch of talk, at a time when the participants were comfortable in front of the camera, of about twenty to forty minutes from the total length of each talk, was then selected for transcription. For the present study, the speech and the gestural data are from eight conversational extracts for a total length of about 160 minutes of talk.

For an understanding of a complete course of action for the construction of meaning across speakers, the domain of analysis is 'a stretch of talk' that comprises the beginning till the end of an exchange about the meaning of a referent. For instance, in a conversation from the data about the strange behavior of a friend, there is a stretch of talk about a discussion on the meaning of an ideophone which sounds like [yuyu], as shown in Example 1. The stretch of talk comprises three turns: In Line 1, the speaker, F1, asserts that the friend behaves in a 'yuyu' way. In the next turn in Line 2, the speaker, F2, shows her understanding of what the ideophone means by the use of the new lexical verb *tiàowǔ* 'dance'. The third turn in Line 3 is the end of the discussion, as F1 displays the acceptance of the second speaker's understanding that dancing is the strange behavior of the friend by the use of the agreement marker *duì* 'right' three times without further discussion.

(1) 1 F1: ..xiànzài biànchéng **yuyu** zhèyàngzi
now become IDEOPHONE like this
'Now, (he) became 'yuyu', like this.'
2 F2: ..ránhòu hái hùi tiàowǔ
then also will dance
'(He) will also dance.'
3 F1: ..duì duì duì
right right right
'Right, right, right.'

Gestures occur during speaking in every stretch of talk in the data, such as the one depicting the ideophone in Example 1 (see Fig. 1): At the time [yuyu] is uttered, F1's hands at each side of the body, with fingers together and curled into fists, move slightly up and down alternately.

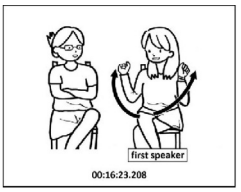
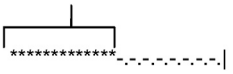


Fig. 1a Gesture for 'yuyu' behaviour



(1) 1 F1: .. xiànzài biànchéng **yuyu** zhèyàngzi
now become IDEOPHONE like this
'Now, (he) became 'yuyu', like this.'

Before the identification of the mimicked gestures, three coders were trained to analyze speech-accompanying gestures. Self-adaptors, like scratching on the arm, despite their occurrence in mimicry (Chartrand and Bargh, 1999), were excluded due to the lack of semantic relation with speech. Coders first learned to identify the boundary of a gesture with respect to three major phases: 'preparation', 'stroke', and 'retraction' (McNeill, 1992). The preparation phase refers to ''the limb mov[ing] away from its rest position to a position in gesture space where stroke begins''; the retraction phase is the ''return of the hand to a rest position''; and in-between is the stroke phase during which ''the meaning of the gesture is expressed'' (McNeill, 1992:83). While both preparation and retraction are optional, the stroke is obligatory. Coders were reminded to pay special attention to

[1] Permissions were obtained from all of the participants to use all of the audio–visual data for research. Moreover, the data from The NCCU Corpus of Spoken Mandarin are a part of an archive of language documentation of the spoken forms of Mandarin, Taiwanese and Hakka in Taiwan (Chui and Lai, 2008). The data can be accessed online at http://spokenchinesecorpus.nccu.edu.tw/. The access date for this study was August, 2013.

gestures which appeared in succession, in that not only is there no retraction and preparation between gestures, but some aspects of meaning could be carried over from one gesture to another. In such a case, analyzing gestures as discrete units is a problem. Furthermore, how can gestures be recognized as semantically related to linguistic expressions or not? Koschmann and LeBaron (2002:263), who analyzed the use of spontaneous gestures in face-to-face interactions in great detail, emphasized ''the importance of including the social and material environment, as well as the conversational history, in the analysis of gesture performance.'' Here, whether or not a gesture has a lexical affiliate and depicts information about it also rests upon the linguistic context and the socio-cultural knowledge about the subject of talk in conversation. The linguistic and socio-cultural knowledge also help separate consecutive gestures. For instance, in a topic about a hornets' nest, the speaker in Example 2 is recalling that a fireman took a ladder and placed it against a tree, climbed up the ladder, and immediately took the nest off from the tree. Two gestures – one for going up a tree (see Fig. 2a) and the other for getting rid of a hornets' nest (see Fig. 2b) – were produced consecutively. First, while uttering the verb *shàngqù* 'go up', the speaker's right hand rises to cheek level with fingers slightly apart and bent to enact the action of going up. Then, at the time the adverbial *zhèyàng* 'like this' is produced, the hand turns clockwise one time to depict the action of taking the nest off from the tree. These two gestures clearly conveyed distinct information about two separate actions.

(2) 1 F2: ...ta bān le tīzi gùoqù
3SG take PRF ladder go across
'He took a ladder there.'

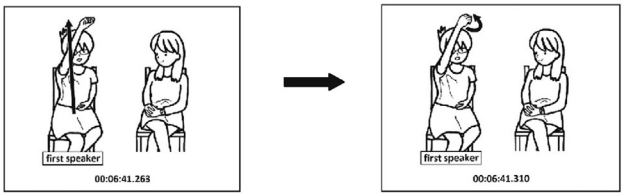


Fig. 2a Gesture for 'going up a tree' Fig. 2b Gesture for 'getting rid of a hornets' nest'

********** | ********** -.-.-.-.-.-.-.-.-.-.-.-.-.-.-.-.-.|
2 .. ránhoù yí **shàngqù** jiù zhèyàng... jiù **zhāixiàlái** le
then once go up just like this just take off PRT
'Then, once (he) went up, (he) just took (it) off like this.'

After the identification of gestural boundaries, the coders were trained to categorize a manual configuration either as an 'iconic gesture', the meaning of which corresponds to the semantic content of the associated speech, a 'metaphoric gesture' representing an abstract idea, a 'deictic gesture' – a pointing at a referent in the immediate speech environment, a 'spatial gesture' designating a gesture space for a referent, or an 'emblematic gesture' having standards of well-formedness (McNeill, 1992; Chui, 2002). Both the identification and the categorization of gesture are necessary for identifying mimicked gestures and for understanding whether certain types of gesture tend to occur in gestural mimicry for the co-construction of meaning. Two coders worked together through two conversational excerpts to identify 20 occurrences of each type of gesture. To establish the reliability of the coding, a third coder independently judged 50% of the gestures that had been identified by the first two coders. The percentage of agreement for the third judgment was 100% for iconic gestures, deictic gestures, spatial gestures and emblematic gestures, and 96% for metaphoric gestures. In cases of discrepancy, the judgment of the first two coders was adopted.

For the study of gestural mimicry, in addition to the categorization of gestures, the coders also learned to analyze gesture features which determine the similarity of gestural forms. In Parrill and Kimbara (2006), the features of motion, hand shape, and location were used to characterize mimicked gestures. Kimbara (2008), on the other hand, focused on one gesture feature – hand shape, and all of the hand shapes in the study could correspond to the ASL counterparts to which they most looked alike. In our data, the determination of such a close resemblance was not straightforward, since most of the spontaneous gestures involve the dynamic movement of fingers, hands and arms. Instead, five gesture features were adopted: 'handedness', 'position', 'orientation', 'hand shape' and 'motion' (McNeill, 1992, 2005) which sufficed to determine

the similarity between gestures in the data in this study. Finally, the coders described the meaning of gestures in words, and looked for lexical affiliates (Schegloff, 1984), if any. Being familiar with all of the criteria, the first two coders worked together and analyzed the gesture features, meaning and lexical affiliate(s) for all of the identified gestures. Then, the third coder analyzed 50% of the data independently. The percentage of agreement for the third judgment was 100% for handedness and orientation, 99.2% for hand shape, and 98.3% for position and motion; the percentage of agreement was 99% for the verbal description of meaning and 97% for lexical affiliation. In cases of discrepancy, the judgment of the first two coders was adopted.

When the training session was finished, the first two coders worked separately to identify stretches of talk and analyze mimicry data from a different dataset for this study. First of all, they looked for the content of the whole of the discussion about meaning across the speakers, such as that about the strange behavior of a friend in Example 1. A total of 63 stretches of talk were found with consensus between the coders. In each stretch, the 'first speaker' brought up a referent, after which the 'second speaker' joined the discussion about its meaning.

The occurrence of gestural repetition was then identified within each stretch of talk. Following Holler and Wilkin's (2011:139) definition that mimicked gestures are ''gestures highly similar in their form *and* in the meaning they depict [italics original],'' the coders understood that 'form' and 'meaning' were the two main criteria to be used in the definition. A high similarity in form is a matter of degree and mimicked gestures could be performed with ''some degree of leeway. . .[or] in a slightly more elliptical form; that is, while the gesture may have looked more sloppy or may have been reduced by a particular semantic aspect, the general conceptualization did not change and a core aspect of the semantic representation was always retained in any gesture coded as mimicked'' (Holler and Wilkin, 2011:140). The high similarity between two gestures in this study was mainly determined by rates of the congruence in the judgment of the five gesture features. Another criterion has to do with meaning, that is, whether or not a mimicked gesture, in addition to having a high similarity in form, also represents the same referent being depicted by its corresponding gesture in the prior context. The context and the content of the utterances made it clear whether the two similar gestures refer to the same referent or not.

The coders looked across the 63 stretches of talk to find instances of mimicry independently. They judged whether two similar gestural forms shared the five gesture features, and whether the two forms depicted the same referent. The coders reached consensus on twelve instances of mimicked gestures which constituted twelve co-referential gesture pairs. In each pair, the initial gesture was produced by the first speaker. Then, the second speaker mimicked it in the next turn. Consider Example 1 again. During the discussion about the meaning of [yuyu], the first speaker, F1, produced a gesture to depict the ideophone in Line 1 (see Fig. 1a). The gesture was then mimicked by F2 (see Fig. 1b) at the time she provides her understanding of the ideophone by the use of the new lexical verb *tiàowǔ* 'dance' in Line 2.

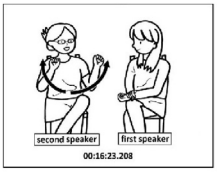


Fig. 1a Gesture for 'yuyu' behaviour Fig. 1b Mimicked gesture for 'yuyu' behaviour

(1) 2 F2: ..ránhòu hái hùi **tiàowǔ**
then also will dance
'(He) will also dance.'

Among the twelve pairs, ten of the mimicked gestures were produced in isolation, and the gestural strokes had clear independent manual configurations that depicted the information coded in speech. The two cases remaining of the twelve pairs included successive gestures; however, although there was no retraction and preparation between the gestures, the strokes depicted distinct kinds of information discretely. One case is the going-up-a-tree gesture followed immediately by the getting-rid-of-a-hornets'-nest gesture in Example 2. The other case is about a friend's swelling condition due to being over-weight, when a swelling-torso gesture comes right after a swelling-face gesture. Both of the gestures, though in succession, involved different body parts. In brief, no matter whether gestures were produced in isolation or in succession, the hand configurations were discrete and depicted independent information in all of the twelve cases.

As to the similarity between the mimicked gestures and their corresponding initial gestures, it was found that the overwhelming majority (92%) differed in one feature at most (five were the same; six differed in one feature); the remaining pair differed in two features. Furthermore, the degree of similarity for each of the five features was also analyzed. The degree of similarity between the two instances in each pair was rated on a five-point scale of agreement: The realization of

Table 1
The congruence rates across the five gesture features.

	Hand shape		Handedness		Position		Motion		Orientation	
*Alike	12	100%	10	83%	11	92%	11	92%	8	67%
Not alike	0	0%	2	17%	1	8%	1	8%	4	33%
Neutral	0	0%	0	0%	0	0%	0	0%	0	0%

* 'Alike': coders (strongly) agreed that the realization of the feature was similar in both gestures; 'Not Alike': coders (strongly) disagreed that the realization was similar in both gestures; 'Neutral': similarity or difference not noticeable.

a feature in both of the gestures of each pair was coded as 'Alike' if the coder chose 'agree' or 'strongly agree', as 'Not Alike' if the judgment was 'disagree' or 'strongly disagree', and as 'Neutral' for the choice 'neutral' on the scale. The judgment was based on the relatively objective spatio-physical manifestation of the features, and coders reached total agreement on their analysis. Table 1 shows the congruence results across the five gesture features, among which total congruence was found for 'hand shape'; and very high congruence was found for 'handedness', 'position' and 'motion'. For the feature of 'handedness', the second speakers in two cases gestured with their left hands, and not with their right hands, as the first speakers had done. The only case of incongruence with respect to 'position' lies in inconsistency in the use of gesture space: The first speaker depicted the size of *gamo* 'a big, flat, round container made from bamboo' on the right and left periphery, whereas the second speaker did so on the lower right and left periphery. 'Motion', which is the movement *per se*, is another feature for which high congruence was found. An exception occurred when the extent of the movement of the hand was not identical between the two corresponding gestures in the pair for *bǎntiáo* 'a kind of rice noodle the dough of which is made into thin sheets, which are then cut into noodles': The first speaker, in addition to putting the right palm flat above the left flat palm to depict the flatness of a layer of rice noodles, also moved the right hand rightward horizontally to signify the layer of the flat noodles; the second speaker repeated only the right-palm-flat-above-left-flat-palm hand shape without moving rightward for the depiction of the layer. Finally, the feature with the lowest level of congruence is 'orientation', suggesting that the direction of the hands and fingers is least consistent in mimicry across speakers. For instance, to gesture *yuèqín* – a musical instrument which has a body with a round back and a flat top, a long neck and strings which are played with the fingers, the fingers of each of the hands of the first speaker were curled into fists, facing each other in front of her chest (see Fig. 5a). The second speaker, though forming two fists in the same way, had the left hand rest on the arm of a sofa instead (see Fig. 5b). In considering the five features together, the deviance in the hand/finger orientation, but the high consistency in the other four features did not affect the conclusion of the analysis that the two gestures were highly similar gestures for the same referent. In summary, while repetitions of gesture are never exact copies, the mimicked data found in the conversational data maintain a very high similarity with their corresponding counterparts.

As to the categorization of the mimicked gestures, total agreement was also reached between the two coders: Five were iconic gestures depicting concrete entities and actions; seven were metaphoric gestures representing abstract ideas, qualities and location. No other types of gesture were found in the dataset. Whether iconic and metaphoric gestures, which convey substantive semantic information, tend to occur in gestural mimicry for the co-construction of meaning, rather than for other types of joint actions is an open question. In summary, the mimicked gestures were co-referential with their respective initial occurrences as associated with lexical constituents in prior turns, and both maintained a high similarity across the five gesture features. The occurrence of mimicked gestures reveals different ways the linguistic and the gestural modalities are involved in establishing meaning across speakers, which will be discussed in the next section.

3. Mimicked gestures and the joint construction of meaning

In Holler and Wilkin's (2011) study where a communication task was set and which required two participants to focus their talk on referents ''in order to figure out whether they are talking about the same thing'' (Holler and Wilkin, 2011:136), a total of 113 mimicked gestures were produced. In face-to-face conversation, however, gestural repetition is not frequent. One reason is that speakers perform many actions other than only co-constructing meaning. Furthermore, speakers do not necessarily mimic others' gestures while they present semantic information about the same referent. Nevertheless, the quantitative data and statistical evidence in Holler and Wilkin (2011) demonstrate that the use of mimicked gestures is by no means a matter of chance. They are a resource for the construction of meaning. How this manual resource is used along with speech in the joint action in conversational discourse is discussed here. This section presents evidence that a complete course of action for the joint construction of meaning reveals the circumstances under which the need to co-construct meaning arises and the way that mimicked gestures are used to accomplish the joint action.

A complete course of action for the joint construction of meaning consists of three phases: the 'presentation phase' in which a meaning is conveyed, the 'collaboration phase' in which the joint action starts and gestural repetition occurs, and the 'completion phase', in which acceptance of the new meaning is acknowledged and the joint action ends. In differing from the two phases as proposed by Clark and his colleagues (Clark and Wilkes-Gibbs, 1986; Clark and Schaefer, 1987;

Clark and Brennan, 1991; Clark and Krych, 2004), the presentation phase here not only conveys a meaning, but also unfolds a contextual situation that will lead to the initiation of the joint action; and the completion phase not only includes the acknowledgment of the new meaning, but also indicates the end of the joint action; and, moreover, the present study further proposes a stage between the presentation phase and the completion phase – the ‘collaboration phase’, during which the joint action is carried out, and it is also during this phase that gestural repetition occurs. Furthermore, the way that mimicked gestures and speech work together to create meaning in daily conversation will be shown to differ from the use of mimicked gestures found in the previous task-based studies. The three phases will be discussed accordingly.

3.1. The presentation phase

In the past task-based studies, participants had to engage in the joint establishment of a referent for mutual understanding. In daily conversation, however, the need for collaboration could still arise when participants are performing other actions. For instance, the speaker, F1, in the first turn in Example 1, is to utter an assessment to characterize the behavior of a friend. It is then necessary to investigate under what kinds of circumstances the joint action might occur, in order to understand the use of mimicked gestures in a more complete way.

In the presentation phase when a referent is brought into discourse, what kinds of contextual situation would initiate the joint establishment of meaning in the next turn? Four kinds of situation were identified in the conversational data by taking into account the interaction between the first and the second speaker in the stretch of talk: ‘difficulty in verbalization’ on the part of the first speaker, ‘lack of clarity’ on the part of the first speaker, ‘disagreement’ on the part of the second speaker, and ‘alignment’ on the part of the second speaker.

3.1.1. Difficulty in verbalization

When the first speaker encounters difficulty in speaking, one consequence is that the turn-construction unit is incomplete and the meaning is not fully conveyed. Then, the need for the co-establishment of meaning readily arises. In Example 3, the first speaker, F1, is making a general statement that if a person were someone with whom she had failed to establish a close relationship, she would idealize the person. A gesture is produced to depict the idea of *lǐxiǎnghuà* ‘idealization’ (see Fig. 3a): The speaker’s right hand rises to cheek level with fingers slightly apart and bent at the time the conjunctive *ránhoù* ‘then’ is uttered, after which the hand turns around clockwise while *lǐxiǎnghuà* is produced. After the assertion, F1 then makes a further attempt to explicate her understanding of ‘idealization’ yet fails to complete the expression of her thought after her utterance of the second degree adverb *hěn* ‘very’ (Line 1). The failure legitimately induces the other speaker to provide a meaning for the idea of idealization in Line 2.

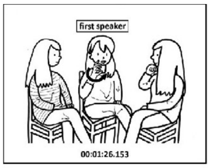


Fig. 3a Gesture for idealization

|~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~***********

(3) 1 F1: ..nǐ débúdào de dōngxī..ránhòu yìzhí dùi tā hěn **lǐxiǎnghuà** de.. hěn
2SG NEG.get DE thing then continuously to 3SG very idealize PRT very
‘For the things that you can’t get, then (you) very much keep idealizing him, very’

2 F2: ... nǐ bǎ tā měihuà
2SG BA 3SG beautify
‘You beautify him.’

3.1.2. Lack of clarity

Another situation where the need for the joint creation of meaning also arises is when there is a lack of clarity in the meaning of the first speaker as s/he introduces a referent in his or her turn. This most commonly happens when new referents are expressed by demonstratives, non-conventional ideophones, or homonyms. The occurrence of the ideophone [yuyu] in Example 1 lacks semantic clarity when it is first presented in Line 1. Example 4 illustrates the presentation of a referent in the form of a demonstrative during a discussion about the shape of the body of a friend, when M2 mentions that the girl was fat at the time that he saw her in a cram school, but that she had become very thin when he saw her again later in the school library

in college. As M1 in Line 1 tries to provide a reason for the change in the girl, his utterance includes the demonstrative *nàge* 'that' as the main predicate and he simultaneously produces a gesture (see Fig. 4a) ''to make more specific the meaning of something that is being said in words'' (Kendon, 2004:176): M1's right hand first rises up with the fingers open from the thigh to the front of the chest at the moment of saying the adverbial *jìu* 'then', and then the left hand also starts rising up with the fingers open. Both hands move up and down alternately four times from the left to the right, depicting the idea of the changes in the shape of the body. The absence of an explicit lexical meaning for the new gestural referent in the prior context prompts M2 to participate in providing a new meaning for the demonstrative (Line 2).

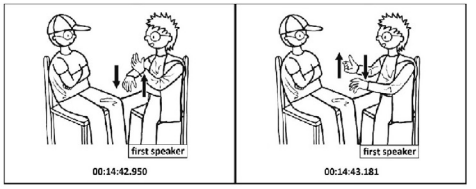


Fig. 4a Gesture for adjustment of the shape of a body

**********-.-.-.|

(4) 1 M1: ..qíshí tā bĕnlái jiù **mán** huì **nàge** a
in fact 3SG originally then quite can that PRT
'Actually, she was originally quite good at that.'

2 M2: ..hái mán néng tiáojié jiù duì le
still quite can adjust then right PRT
'(She) is quite good at adjusting (her body shape).'

3 M1: duì a tā ..tā yĕ pàng shòu pàng shòu pàng shòu.. zhèyàng
right PRT 3SG 3SG also fat thin fat thin fat thin like this
'Right. She…she's fat, and then thin, and then fat, and then thin, and then fat, and then thin, like this.'

3.1.3. Disagreement

The third situation has to do with disagreement, in that the second speaker may engage in the joint establishment of meaning because s/he does not agree with what the first speaker has uttered in the prior turn. The talk in Example 5 is about the kind of musical instrument that is played by a character in a movie. The first speaker, F, uses a general term *yuèqí* 'musical instrument' in speech (Line 1) but gestures the particular kind of instrument that requires the use of a bow to play (Fig. 5a): During the pause after the classifier *zhŏng* 'kind', the speaker's right hand goes up to shoulder level with the fingers curled into a fist as if holding a bow; the left hand rises to waist level, also with fingers curled into a fist as if holding the lower part of the instrument. Then, at the time *yuèqí* is uttered, the right hand moves horizontally to the left one time to enact the playing of a string instrument that requires the use of a bow. Since M as the second speaker holds a contrary opinion about the referent *yuèqí* in regard to the instrument played in the movie, as evidenced by the negative word *méiyŏu* 'no' at the beginning of his turn in Line 2, he then brings up a different understanding in his turn.

Fig. 5a Gesture for musical instrument with a bow

|~~**********-.-.|

(5) 1 F: .. màobó shì nà zhŏng... zhuānyè de nà zhŏng... **yuèqì** de ..nĭ zhīdào ma.. suŏyĭ
Maobo COP that kind professional DE that kind musical instrument DE 2SG know QST so
'Maobo (used) that kind of professional musical instrument, you know. So'

2 M: (0) méiyǒu.. tā shì nàge...(1.1)[<L5 jù xuánzǎi L5>
NEG 3SG COP that play plucked lute with a wooden body
'No, he played a kind of plucked lute with a wooden body.'

3 F: .. duì a duì a... jiù shì nà zhǒng... yuèqì de
right PRT right PRT EMP COP that kind musical instrument PRT
'Righ, right. It's that kind of musical instrument.'

3.1.4. Alignment

Finally, the second speaker may participate in the joint establishment of meaning even though s/he does not object to what the first speaker talks about in the prior turn. In Example 6, the conversational topic is about feeling itchy during the harvesting of crops in a field. In Line 1, the first speaker, M1, is explaining that there are aristae *máng* on the stems. At the same time, a gesture is produced for aristae (Fig. 6a): While uttering the causative *yīnwèi* 'because', M1's right hand rises to shoulder level; his left hand goes to chest level. At the time the nominal *máng* is produced, his right hand curves into the palm, whereas his left-hand fingers come together. The configuration as a whole enacts the holding the stem of a crop on which there are aristae. Another participant M2 supports the idea by providing more information about the same gestural referent *máng* in the next turn (Line 2).

Fig. 6a Gesture for aristae

|~~~~~~~~~~~~~~~~~*********

(6) 1 M1: .. yīnwèi tā yǒu **máng** a
because 3SG have arista PRT
'Because they've got aristae.'

2 M2: ..shàngmiàn yǒu nàge háomáo
on have that fine hair
'(They've got) those fine hairs (on them).'

Among the four types of contextual situation mentioned above, the joint action most likely ensues in the first three types when meaning in the prior utterance is incomplete or lacks clarity, or when the second speaker has a different understanding of the meaning. Nonetheless, in five cases, the second speaker still provides a meaning for the same referent even when s/he does not disagree to what the first speaker has uttered about it. Participation in the joint construction of meaning could be a way to confirm the understanding of a referent between the participants. No matter whether the first speakers' utterances are complete or not, clearly or vaguely expressed, recognized or not, they are assertions or assessments about the quality of states, activities, or processes, or about the characteristics of people or objects. The utterances each include a referent in gesture. Among all the twelve gestures, most depict abstract ideas (seven instances) such as 'idealization' in Example 3; the remaining instances depict concrete entities or activities such as the gesture for the playing of musical instruments in Example 5. In considering all of the gestural referents together, the overwhelming majority (11 out of the total of 12) carry new information, as the corresponding lexical expressions have not been brought up in the preceding context. Old information tends not to be represented manually (McNeill and Levy, 1993; Chui, 2005), yet exceptions occur when the same referent is mentioned in different topics. In the data, the only exception is for *bǎntiáo* 'rice noodles'. The lexical referent was first mentioned in a conversational topic about different kinds of food for a class reunion. When it was brought up a second time in the conversation, the talk rather focused on the features of the kind of rice noodles being described, during which a gesture was produced for that kind of noodles: The right palm is put flat above the left one to depict the flatness of a layer of rice noodles.

In summary, the findings demonstrate that in the presentation phase, a variety of situations other than that of communicating a referent for mutual understanding would prompt the other speaker to participate in the co-construction of meaning for a new gestural referent which has just been brought up in the prior assertion or assessment, be it concrete or abstract, or be it lexical or not.

3.2. The collaboration phase

After the presentation stage is the collaboration phase during which the other participant expresses his or her own understanding of a referent being mentioned in the first speaker's turn. The co-construction of meaning starts and proceeds on the part of the second speaker. This is also the phase during which mimicked gestures were found to occur. In Holler and Wilkin (2011), they proposed that gestures of this kind fulfill the functions of 'presentation', 'acceptance', and 'displaying incremental understanding':

> The presentation category was loosely based on Clark and Wilkes-Gibbs' (1986) definition of the term and comprised all noun phrases as well as more complex descriptions of the stimuli. . .. The acceptance category includes all those cases that, based on Clark and Wilkes-Gibbs, fulfilled the function of asserting acceptance of a presentation. The displaying incremental understanding category refers to cases where interlocutors signaled that they had understood to some degree but were still trying to figure out exactly which one the intended referent is. . ., or by performing the gesture slightly more hesitantly while examining their cards after looking at their interactant (Holler and Wilkin, 2011:141).

However, it will be shown in this section that there are different ways the linguistic and the gestural modalities are involved in establishing meaning for the same referent across speakers in conversation. First, the ways in which the second speakers respond to the various preceding contextual situations as mentioned in section 3.1. are discussed.

3.2.1. Difficulty in verbalization

In Example 3, as the first speaker, F1, fails to finish elucidating her idea about *lǐxiǎnghuà* 'idealization', the second speaker, F2, does not finish the previous speaker's incomplete utterance. Instead, she provides her own understanding of what may be meant by the idealization of a person by formulating a new statement with a different verb *měihuà* 'beautification' (Line 2). In gesture, rather than depicting *měihuà*, F2, simultaneous with the verb, uses a metaphoric gesture mimicking the gesture for *lǐxiǎnghuà* which was produced in the presentation phase: Just as in the enactment of the initial gesture (Fig. 3a), the second speaker turns her right hand at the side of her right face clockwise with the same hand shape during the collaboration phase. See Fig. 3b.

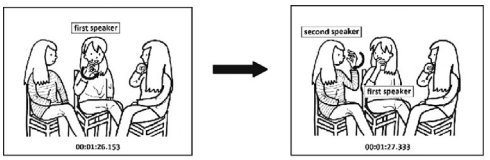


Fig. 3a Gesture for idealization Fig. 3b Mimicked gesture for idealization

					*******		
2	F2:	... nǐ	bǎ	tā	**měihuà**	...	duì
		2SG	BA	3SG	beautify		right

'You beautify him.'

3.2.2. Lack of clarity

The need for the joint creation of meaning also arises when a new referent in the presentation phase is expressed by a demonstrative or an ideophone which lacks explicit meaning, or by a homonym or a synonym with semantic ambiguity. Disambiguation by means of repeating a prior gesture can be found in Kimbara (2006). In this study, a lack of semantic clarity can also be resolved by gestural repetition. As illustrated in Example 4, the first speaker's use of *nàge* 'that' is accompanied by a gesture depicting changes in the shape of a body (Fig. 4a). To provide an explicit meaning for the demonstrative, the second speaker, M2, starting from the same adverbial *mán* 'quite' in his turn in Line 2 in Example 4,

mimics the body-shape-adjustment gesture (Fig. 4b), by copying the handshape of M1's gesture in front of the chest, its handedness (both hands) and the direction of movement (up and down).

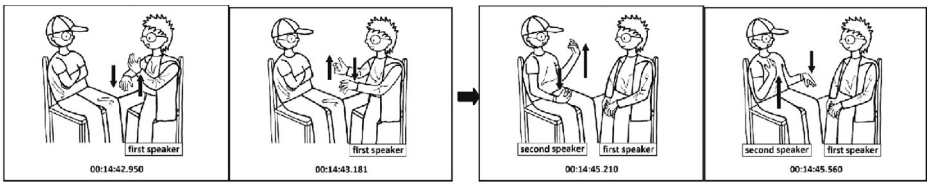


Fig. 4a Gesture for adjustment of the shape of a body

Fig. 4b Mimicked gesture for adjustment of the shape of a body

2 M2: ..hái mán néng **tiáojié** jiù duì le
still quite can adjust then right PRT
'(She) is quite good at adjusting her body.'

3.2.3. Disagreement

To express an understanding contrary to what the first speaker has said about a referent, the second speaker can also produce a mimicked gesture while constructing more meaning for the same referent. In Example 5 about musical instruments, the second speaker disagrees with the statement of the first speaker that the instrument the character in the movie plays requires a bow. Rather, he mentions that it is the type that is played with the fingers, as represented in speech by *yuèqìn* 'plucked lute with a wooden body' (Line 2). But, what is of note is that in gesture, instead of enacting *yuèqìn*, which is played with the fingers, the speaker mimics the first speaker's gesture (Fig. 5a and b), as the fingers of M's right hand form a fist like holding a bow, and the right hand moves in the same leftward direction to enact the idea of playing music with a bow. This visible evidence demonstrates that the speaker is considering *yuèqì* in gesture and *yuèqìn* in speech at the same time.

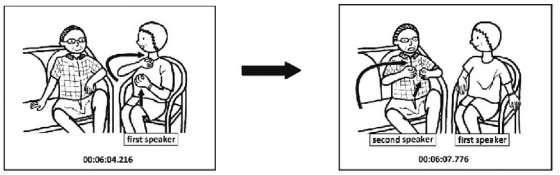


Fig. 5a Gesture for musical instrument with a bow

Fig. 5b Mimicked gesture for musical instrument with a bow

2 M: (0) méiyǒu.. tā shì nàge...[<L5 jù **xuánzǎi** L5>
NEG 3SG COP that play plucked lute with a wooden body
'No, he played a kind of plucked lute with a wooden body.'

3.2.4. Alignment

Finally, without opposition to what the first speaker has uttered, additional meaning can still be established for the referent. In the topic about feeling itchy during the harvesting of crops in Example 6, the second speaker, to show alignment with the first speaker about the gestural referent *máng* 'arista', provides additional characterization of the crops – having *háomáo* 'fine hair' on the stems (Line 2) and also repeats the initial gesture (see Figs. 6a and b), mimicking the hand shape of holding a stem with both hands in front of the chest, as well as the orientation of hands and fingers being together and curved in.

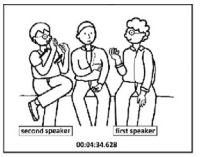


Fig. 6a Gesture for aristae

Fig. 6b Mimicked gesture for aristae

****

2 M2: ..shàngmiàn yǒu nàge **háomáo**
on have that fine hair
'(They've got) those fine hairs (on them).'

The four instances of mimicked gestures discussed above manifest one of the ways in which speech and gesture collaborate to create and provide meaning for a referent, be the action prompted by the first speaker's experiencing an obstacle to verbalization or by a lack of clarity in the utterance, or by the second speaker's alignment or disagreement with what the first speaker has uttered. The joint action proceeds during the construction of utterance, with the second speaker mimicking an initial gesture without repeating its corresponding lexical constituent. Instead, the mimicked gesture is mostly performed in synchronization with the production of a new linguistic expression which conveys new information about the initial gestural referent.

In some instances, linguistic and gestural repetition both occur; still, new information about the same referent is presented by a new linguistic expression in the utterance. The talk in Example 7 focuses on a location of Taipei – Tianmu. The first speaker, M, in Line 1 regards Tianmu as a place where there is a park near the Donghu metro station. The locative demonstrative *nàlǐ* 'there' refers to Tianmu, which is also depicted by a metaphoric gesture which conceptualizes the place as an entity with boundaries (see Fig. 7a): At the time that the demonstrative is uttered, both of M's hands rise to chest level and are held apart with the palms facing one another and the fingers are slightly curled, enacting a bounded area for Tianmu. The second speaker, F, expresses disagreement by first using a co-referential demonstrative *nēige* as the subject in Line 2 and by mimicking M's gesture (Fig. 7b) with the same hand shape and finger orientation at chest level. Then, the second speaker asserts that the area M refers to was Tianmu in the past. For the two participants to jointly create a referent for Tianmu is thus accomplished by linguistic and gestural repetition.

Fig. 7a Gesture for Tianmu

****

(7) 1 M: ..donghú nàge.. yí ge shéme... gōngyuán nàbiān.. jiùshì.. **nàlǐ** cái [shì]
Donghu that one CL whatever park there that is there just COP
'(As to) Donghu, there's a…whatever…park there, that is, it's there.'

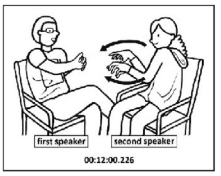


Fig. 7b Mimicked gesture for Tianmu

~~~~~~~~~~~~~~~~~~~~~~********

2 F: ..[nēi ge] shì yǐqián de **tiānmǔ**
that COP before DE Tianmu
'That was Tianmu in the past.'
~~~~~~~~~~~~~~~~~~~~~~

There is a second way for collaboration between speech and gesture in a discussion about size, where the new meaning is conveyed exclusively by the gestural modality. In Example 8, both the first and the second speakers present their own assessment of the size of hornets by uttering the same pro-form *zhèyàng* 'like this' across turns. Exactly how big the hornets are is conveyed in gesture exclusively (see Fig. 8a): At the time the first speaker, F1, utters *dàgài* 'approximately' in Line 1, the right thumb and index finger at chest level fully extend and draw apart, as the remaining fingers are curled into the palm; the space between them represents the size of the insect. Since the second speaker, F2, does not agree with the size, she repeats the first speaker's gesture yet with her left hand rather than the right hand, as the first speaker has done. She also reiterates the associated word *zhèyàng* (Line 4) without further elaboration in words, and new information is depicted by the slight modification of the mimicked gesture by moving the thumb and index finger a bit further apart to enact a larger size than the initial counterpart (Fig. 8b). The first speaker also disagrees and gives assurance of the smaller size again in Line 5 by moving her gesture forward to a shared central space. Then, as the first speaker's and the second speaker's gestures meet in the shared space, the second speaker's larger-size gesture becomes smaller to match the size represented in the first speaker's gesture (Fig. 8c). Both speakers then hold the same gesture as they discuss further about the size of hornets in Lines 6 and 7.

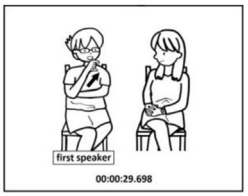


Fig. 8a Gesture for the size of hornets

(8) 1 F1: nàge f- huángfēng hěn dà zhī ye…dàgài **zhèyàng**
that REPAIR hornet very big CL PRT approximately like this
'The…hornets were very big, approximately like this.'

2 F2: …pì la
nonsense PRT
'Nonsense.'

****

3 F1: …zhēnde zhēnde.. [chābùduō zhèyàng]
really really almost like this
'Really, really, (they're) almost like this.'

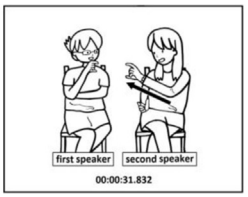


Fig. 8b Mimicked gesture for the size of hornets

****

4 F2: [**zhèyàng**]
like this
'(They're) like this.'

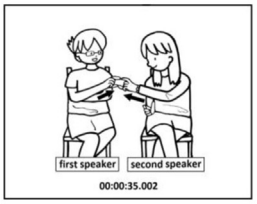


Fig. 8c Mimicked gesture for the size of hornets

****

5 F1: ..méiyǒu la zhèyàng la zhèyàng
NEG PRT like this PRT like this
'No, (they're) like this, like this.'

****

6 F2: ..nǐ shì shuō chuán cháng jiù jiā shǒu jiā jiǎo ma
2SG COP say all length just add hand add foot QST
'Do you mean the whole length including hands and feet?'

****

7 F1: (0)duì duì duì
right right right
'Right, right, right.'

((THE FOUR FOLLOWING TURNS ARE JOKES ABOUT HOW THEIR GESTURES LOOK LIKE A FORM OF ALIEN COMMUNICATION))

****

8 F1: …jiù dàgài shì zhèyàng
then approximately COP like this
'Then, (they're) approximately like this'

Example 8 also provides visible evidence that the joint construction of meaning is an incremental process, with the taking of multiple turns to establish the meaning of a referent at issue before consensus is reached. The collaboration phase of the joint action in the example consists of six turns (Line 2 to Line 7) for the discussion between the speakers about the size of hornets. During the collaboration phase, the initial size gesture occurs in F1's turn in Line 1 and the stroke holds till the end of the phase in Line 8. The second speaker first shows disagreement in speech in Line 2, and then, by producing a mimicked gesture in her second turn (Line 4), she slightly draws her thumb and index finger apart for a larger size. With no consensus, the first speaker in Line 5 has her smaller-size gesture meet the second speaker's to make it smaller. Then, the co-establishment of meaning continues as the second speaker, in Line 6, holds the smaller-size gesture and raises a question in speech about whether to include the length of the legs of a hornet in calculating the size of hornets, after which the first speaker provides a positive response in Line 7. Agreement comes in Line 8, when the first speaker again gives assurance of the smaller size without any objection from the second speaker.

Given the two ways in which mimicked gestures are used along with speech, our data demonstrate that it is more frequent to represent new meaning in speech while the first speaker's gesture is also mimicked. Whether speakers commonly express new meaning only by the use of a mimicked gesture with slight modification for certain types of semantic information such as 'size' awaits future research (Beattie and Shovelton, 2006).

3.3. The completion phase

The last stage of the joint action is the completion phase during which the new meaning is recognized and accepted and the joint action ends. The discourse marker *duì* 'right' in Mandarin, which functions to convey agreement (Chui, 2002; Wang et al., 2010), and *o* 'I see' are commonly used in the acceptance phase to acknowledge the newly-established meaning, as illustrated in Line 3 in Examples 1, 4 and 5. Acceptance can also be signaled non-verbally, for instance by head nods. The use of the same mimicked gesture also indicates acceptance, just as in the case of F2's gestural action in Example 8 in which F2 accepts F1's assurance in Line 7 about the smaller size of the hornets by mimicking F1's smaller-size gesture without any speech. Moreover, further elaboration in words about the gestural referent after the agreement marker indicates further confirmation of the new meaning provided by the second speaker. In Example 5 about musical instruments, after the second speaker has mimicked the plucking gesture which forms the basis for establishing the correct type of the musical instrument which is played by the character in the movie, which in fact is a string instrument that is played by the hands plucking some strings, rather than a bow being drawn across them, the first speaker, F, in her next turn (Line 3 below) agrees on the new referent by uttering the agreement marker *duì* two times, followed by an affirmation of what M has mentioned about the musical instrument.

(5)	3	F:	**..duì**	a	**duì**	a...	jiù	shì	nà	zǒng...	yuèqì		de
			right	PRT	right	PRT	EMP	COP	that	kind	musical	instrument	PRT
			'Right, right. It's that kind of musical instrument.'										

Finally, the lack of explicit acceptance also occurs, as the talk moves on to other subject matter after the completion phase. In Example 3, the discussion of meaning only takes two turns. After the second speaker has provided a new meaning for *lǐxiǎnghuà* 'idealization', the addressee, F1, who is also the first speaker, does not signal acceptance verbally or non-verbally; instead, the talk moves on without opposition. The absence of objection in the completion phase typically implicates agreement and mutual understanding, and the joint action is achieved.

4. Discussion and conclusion

For the participants to discuss meaning together is a common type of joint action in daily conversation. Despite the fact that the joint creation of meaning across speakers is readily accomplished through the collaboration of speech and gesture in sequential context, the establishment of meaning for the same referent is not always achieved by the use of mimicked gestures along with speech. It can be accomplished by a single modality or both modalities. When speech and gesture are used, different speakers can choose to gesture different semantic aspects of the same referent, given that ''[r]eferents tend to have more than one encoding possibility in gesture'' (Kimbara, 2008:123), and mimicked representational gestures are not conventionalized. Thus, the referents can be depicted in a wide variety of ways. In our data, a speaker, in characterizing knee braces, moves both hands, with all fingers spread and slightly bent, toward the right knee and touches both sides of the knees. The gesture signifies the particular body part for wearing of knee braces, which corresponds to the utterance *yīge* 'one' *xīgài* 'knee' *de* 'POSS' *nàge* 'that'. . .*zhuāng* 'put' *zhèlǐ* 'here' 'that is for knees. . .you put it here (i.e., on the knee)'. A second speaker in the next turn jointly establishes meaning for the same

referent and provides more information about knee braces. In gesture, she rather depicts knee braces with regard to their length: While saying *yīge* 'one' *hěn* 'very' *cháng* 'long' *hěn* 'very' *cháng* 'long' *de* 'POSS' *nàge* 'that', 'that is very long, very long', the speaker moves both hands from the top of her left thigh down to the knee for the expression of the length.

The second speakers in our data repeat the first speakers' gesture, rather than depicting different semantic aspects of the same referent, in the joint construction of meaning. In the past, neither the course of the joint action nor the use of linguistic-gestural resources to accomplish the action was fully examined in conversational discourse. The current study, based on conversational data, examined gestural mimicry in the whole course of joint creation of meaning. The study found that the complete course of action for the joint construction of meaning consists of three consecutive phases: First, the beginning of the joint action is the presentation phase, when the first speaker provides meaning in a way that would prompt the other participant to engage in the joint action. In the next phase, the collaboration phase, the joint action starts and the second speaker provides new meaning together with a mimicked gesture. The end of the joint action is the completion phase, where the first speaker acknowledges the new meaning and the collaboration ends.

In conversation, the use of mimicked gestures in the collaboration phase was found to be different from the gestural repetition found in other situations. Two ways to present new meaning about the same referent across speakers by the use of linguistic and gestural resources were found. In conversation, one way is to present new meaning about the same referent mainly by gestural repetition with slight modification. Certain types of semantic information, such as shape, size and spatial relations, have been claimed to be more readily expressed in gesture than in speech. Beattie and Shovelton's (2006:63) study shows that ''high importance size information was significantly more likely to be encoded in gesture rather than in speech, whereas low importance size information was more likely to be encoded in speech rather than in gesture.'' This is borne out by the size referent in Example 8, in that information about size, the focus of the discussion, is encoded only in gesture.

The other way to present new meaning about the same referent is with speech and gesture, in that the second speaker conveys a new meaning with a new lexical expression, and at the same time mimics the first speaker's gesture without encoding of the new information. What is of note is that the second speaker would rather mimic the previous speaker's gesture than produce a different one for the new constituent, even though the speaker disagrees with the meaning or referent being uttered and/or depicted, as exemplified in Example 5. The mimicked gesture functions as a semantic foundation that is shared by the two speakers across turns. With this semantic foundation, further meaning can be established for the referent at issue.

When both the linguistic and gestural modalities are involved in the joint construction of meaning, the pairing of the initial gesture and its mimicked counterpart and their collaboration with speech indicates the use of a cross-modal strategy to accomplish the joint construction of meaning in daily communication. Given the *principle of least joint effort* (Clark and Wilkes-Gibbs, 1986; Clark and Schaefer, 1987; Clark and Brennan, 1991; Clark, 1996), the provision of new meaning on a shared foundation through the use of cross-modal resources can be efficient to facilitate the simultaneous realization of the expression of shared knowledge in gesture and new meaning in speech within a clausal unit. In short, together with the findings in previous studies, the study of the use of mimicked gestures in conversational discourse provides a deeper understanding of the occurrence of gestural repetition in different types of interaction.

The findings of the study also help understand the role of gesture in communication. Numerous studies have already demonstrated that gestures express a large amount of information not represented in speech (see the review in Holler and Wilkin, 2011). The present study, together with the other studies on mimicked gestures, provides further evidence in support of the view that spontaneous gestures are communicative. Examples (3) and (4) in Kimbara (2006) show that mimicked gestures were produced when participants re-formulated speech and gesture or co-constructed cartoon content. In this study, participants' observation of others' gestures was pertinent during the formation of joint actions, in that the production of the same gesture shows that participants are attentive listeners (de Fornel, 1992), and the use of mimicked gestures shows strong involvement in the interaction (de Fornel, 1992; Tabensky, 2001), demonstrating that the second speakers pay attention to the gesture of the first speaker in the previous turn, decode the gestural meanings, and formulate the next utterances by the incorporation of information from the manual modality. In the next utterance, a new aspect of the meaning of the referent under discussion is given on the basis of a mutual semantic foundation as depicted by the mimicked gesture. In brief, the re-enactment of a previous gesture manifests that the first gesture has already been recognized, understood, and integrated into a new turn-construction unit.

In conversation, mimicked gestures are also used in other types of joint actions, such as the co-construction of TCUs or the co-enumeration of objects. They can also be used to display attentiveness, understanding, and support without speech. Moreover, gestures of this type could be produced for the performance of impolite acts (Culpeper, 2011a, 2011b). In the future, it is hoped that more conversational data and more mimicked gestures will be available for the study of their social-affective functions and of their manifestation in the socio-cultural aspects of language use.

Appendix A. Speech and gesture transcription conventions

Transcription of speech

[ ]	speech overlap
...(N)	long pause
...	medium pause
..	short pause
<L5 L5>	switching from Mandarin to Southern Min

Transcription of gesture

Kendon's (2004) transcription conventions for gesture were adopted.

|	gesture phrase boundaries
~ ~ ~	preparation phase
~_~_~	pre-stroke hold
****	stroke phase
****	post-stroke hold
-.-.-	recovery phase
/	gesture phase boundaries

The time code shown at the bottom of each panel in the drawing frames in examples is expressed in *hours: minutes: seconds.milliseconds*.

Appendix B. Abbreviations of linguistic terms

2SG	second person singular
3SG	third person singular
BA	morpheme *ba*
BC	backchannel
CL	classifier
COP	copula verb
DE	morpheme de
EMP	emphatic morpheme
IDEOPHONE	ideophone
NEG	negative morpheme
POSS	possessive morpheme
PRF	perfective morpheme
PRT	discourse particle
QST	question morpheme
REPAIR	repair element

References

Beattie, Geoffrey W., Shovelton, Heather K., 2006. When size really matters: how a single semantic feature is represented in the speech and gesture modalities. Gesture 25, 63–84.

Chartrand, Tanya L., Bargh, John A., 1999. The chameleon effect: the perception-behavior link and social interaction. Journal of Personality and Social Psychology 76 (6), 893–910.

Chui, Kawai, 2002. Categorization of gestures in communication. In: Su, I.-W., Lien, C., Chui, K. (Eds.), Form and Function: Linguistic Studies in Honor of Shuanfan Huang. Crane Publishing, Taipei, pp. 105–129.

Chui, Kawai, 2005. Topicality and gesture in Chinese conversational discourse. Language and Linguistics 6 (4), 635–654.

Chui, Kawai, Lai, Hueiling, 2008. The NCCU corpus of spoken Chinese: Mandarin, Hakka, and Southern Min. Taiwan Journal of Linguistics 6 (2), 119–144.

Clark, Herbert H., 1996. Using Language. Cambridge University Press, Cambridge.

Clark, Herbert H., Brennan, Susan E., 1991. Grounding in communication. In: Resnick, L.B., Levine, J.M., Teasley, S.D. (Eds.), Perspectives on Socially Shared Cognition. American Psychological Association, Washington, DC, pp. 127–149.

Clark, Herbert H., Krych, Meredyth A., 2004. Speaking while monitoring addressees for understanding. Journal of Memory and Language 501, 62–81.

Clark, Herbert H., Schaefer, Edward F., 1987. Concealing one's meaning from overhearers. Journal of Memory and Language 26 (2), 209–225.

Clark, Herbert H., Wilkes-Gibbs, Deanna, 1986. Referring as a collaborative process. Cognition 22 (1), 1–39.

Culpeper, J., 2011a. Impoliteness: Using Language to Cause Offence. Cambridge University Press, Cambridge.
Culpeper, J., 2011b. Politeness and impoliteness. In: Andersen, G., Aijmer, K. (Eds.), Pragmatics of Society. Mouton de Gruyter, Berlin, pp. 391–438.
de Fornel, Michel, 1992. The return gesture: some remarks on context, inference and iconic gesture. In: Auer, P., Luzio, A.D. (Eds.), The Contexualisation of Language. John Benjamins, Amsterdam, pp. 159–176.
Goldin-Meadow, Susan, 1999. The role of gesture in communication and thinking. Trends in Cognitive Science 3, 419–429.
Goodwin, Charles, Goodwin, Marjorie H., 1992. Context activity and participation. In: Auer, P., Luzio, A.D. (Eds.), The Contexualisation of Language. John Benjamins, Amsterdam, pp. 77–99.
Holler, Judith, Wilkin, Katie, 2011. Co-speech gesture mimicry in the process of collaborative referring during face-to-face dialogue. Journal of Nonverbal Behavior 35, 133–153.
Kendon, Adam, 2004. Gesture: Visible Action as Utterance. Cambridge University Press, Cambridge.
Kimbara, Irene, 2006. On gestural mimicry. Gesture 6 (1), 39–61.
Kimbara, Irene, 2008. Gesture form convergence in joint description. Journal of Nonverbal Behavior 32 (2), 123–131.
Koschmann, Timothy, LeBaron, Curtis, 2002. Learner articulation as interactional achievement: studying the conversation of gesture. Cognition and Instruction 20, 249–282.
Lerner, Gene H., 2002. Turn-sharing. In: Ford, C.E., Fox, B.A., Thompson, S.A. (Eds.), The Language of Turn and Sequence. Oxford University Press, Oxford, pp. 225–256.
McNeill, David, 1992. Hand and Mind – What Gestures Reveal about Thought. University of Chicago Press, Chicago.
McNeill, David, 2000. Language and Gesture. Cambridge University Press, Cambridge.
McNeill, David, 2005. Gesture and Thought. University of Chicago Press, Chicago.
McNeill, David, Levy, Elena T., 1993. Cohesion and gesture. Discourse Processes 16 (4), 363–386.
Parrill, Fey, Kimbara, Irene, 2006. Seeing and hearing double: the influence of mimicry in speech and gesture and observers. Journal of Nonverbal Behavior 30 (4), 157–166.
Sacks, Harvey, Schegloff, Emanuel A., Jefferson, Gail, 1974. A simplest systematics for the organization of turn-taking for conversation. Language 50, 696–735.
Schegloff, Emanuel, 1984. On some gestures' relation to talk. In: Atkinson, J.M., Heritage, J. (Eds.), Structures of Social Action: Studies in Conversation Analysis. Cambridge University Press, New York, pp. 266–296.
Stivers, T., Enfield, N.J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., Hoymann, G., Rossano, F., De Ruiter, J.P., Yoon, K.-E., Levinson, S. C., 2009. Universals and cultural variation in turn-taking in conversation. Proceedings of the National Academy of Sciences of the United States of America 106 (26), 10587–10592.
Tabensky, Alexis, 2001. Gesture and speech rephrasings in conversation. Gesture 1 (2), 213–235.
Wang, Yu-Fang, Tsai, Pi-Hua, Goodman, David, Lin, Meng-Ying, 2010. Agreement, acknowledgement, and alignment: the discourse-pragmatic functions of *hao* and *dui* in Taiwan Mandarin conversation. Discourse Studies 12 (2), 241–267.

Kawai Chui is Professor of Linguistics at National Chengchi University (NCCU), Taipei, Taiwan. Her research is mainly concerned with the use of language and gesture in spoken discourse and with the recurrent patterns of expression across the two modalities, to understand the relationship among language, gesture and embodied cognition as manifested in daily communication.