When Words Fail: Collaborative Gestures During Clarification Dialogues

Patrick G. T. Healey and **Nicola Plant**
Queen Mary University of London
Cognitive Science Research Group
London, UK

Christine Howes
University of Gothenburg
Centre for Language Technology
Gothenburg, Sweden

Mary Lavelle
King's College London
Institute of Psychiatry, Psychology and Neuroscience
London, UK

Abstract

Research on co-speech gestures has primarily focussed on speakers. However, in conversation non-speaking addressees can also gesture. Often this is to provide concurrent feedback such as backchannels but sometimes it involves gestures that relate to the specific content of the speaker's turn. We hypothesise that non-speakers should contribute most actively during clarification sequences i.e. at the moments when mutual-understanding is threatened. We test this hypothesis using a corpus of story-telling dialogues, captured using video and motion capture. The results show that during clarification sequences speaker and non-speaker behaviours tend to merge. Non-speakers in particular move their hands faster and produce more than twice as many content-specific specific gestures in overlap with speaker turns. These results underline the collaborative nature of conversation, the strategic importance of non-verbal resources for sustaining mutual-understanding and the critical role of clarification and repair in successful communication.

Introduction

Conversation is a collaborative, interactive activity in which both speakers and addressees concurrently contribute to the production of each turn (Goodwin 1979; Clark 1996). While a significant amount has been written about individual gesture production and comprehension outside the context of conversation less is known about how gestures are deployed during live interaction. What is known suggests that conversational context has significant effects both on what kinds of gesture are actually produced (Bavelas et al. 1992; 1995) and how they are deployed (Özyürek 2002).

Although speaker's gestures are the normal focus of attention (Holler and Wilkin 2011) in conversation non-speaking addressees can also contribute to the interaction. Normally people do not gesticulate when silent (Gullberg 1998), partly because hand movements can be interpreted as a bid for the floor. However, there are occasions when non-speakers do contribute non-verbally (Bavelas et al. 2000; Bavelas, Coates, and Johnson 2006).

One common conversational function of non-speaker's gestures is to provide concurrent feedback to speakers about how addressees are responding to their unfolding contribution. The best-known form of feedback of this kind is concurrent verbal back-channels such as "mmm", "uhuh" or nods (Yngve 1970) that provide evidence that addressees have understood the contribution so far. This is often used, for example, to check that a particular installment of a turn such as the introduction of a new referent has been understood before proceeding (Clark and Wilkes-Gibbs 1986). Bavelas and colleagues demonstrated the importance of both generic non-speech back-channels and addressee's content-specific responses such as wincing or smiling at appropriate moments to the fluency and effectiveness of speaker's turns (Bavelas et al. 2000; Bavelas, Coates, and Johnson 2006).

Conversation also creates opportunities for kinds of closely coordinated gesture use, such as mutual-modification of gestures (Tabensky 2001; Furuyama 2002) and the use of shared topic spaces (Emmorey and Reilly 1995), that are unique to conversation. For example (Furuyama 2002) describes how during instructional dialogues involving descriptions of how to fold origami figures the learners will produced gestures co-ordinated with the instructors speech. These gestures demonstrate understanding of an instruction and sometimes involve movements that are closely co-ordinated with the instructors gestures. As (Furuyama 2002) points out, these non-speaker gestures are especially interesting because most models of gesture production have assumed that the components of the cognitive or computational system that produce them are "mechanically connected" through neural, muscular or skeletal hardware. The integration of gesture and speech across multiple participants presents an obvious challenge to this assumption.

The examples of non-speaker gestures above illustrate how non-speakers can provide positive evidence of understanding for a speaker. However, people also take advantage of concurrent feedback to signal problems in an interaction (Clark and Schaefer 1989) and this can also be in the form of gestures (Gullberg 1998; Holler and Wilkin 2011). This negative feedback can take the mild form of more provisional acceptance or can involve overt attempts to initiate or respond to clarifications. The mechanisms involved in detecting and dealing with problems in conversation in this way have received detailed attention from conversation analysts (Schegloff 1987; 1992) and their approach informs the analysis reported here. Most of the data for the conversation analytic analyses has been qualitative and focused on the verbal aspects of repair although more recent work has highlighted multi-modal aspects of these processes (Mondada 2014).

A key motivation for the analysis presented here is the intuition that the provision of negative evidence of understanding defines critical moments in interaction, requiring people to change trajectory and produce edits or amendments to keep a conversation on track. (Healey 2008) argues that these moments are much more important for the co-ordination of understanding than positive evidence and require active collaboration to resolve. This leads to the general hypothesis that communicative resources should be at a premium at the points in conversation where problems with mutual understanding are encountered. Gestures, especially non-speakers gestures, provide one such resource. They can be used to provide additional information, such as spatial relationships between objects, that are difficult to convey using speech alone and they can be used to make contributions which do not compete with the speech channel.

Previous work has shown that other forms of non-speech feedback, specifically nodding, is correlated with the production of within-turn self-repairs or 'disfluencies' (Healey et al. 2013).[1] Importantly, this effect is more marked for the addressees of a problem turn than for the speaker. This suggests that when a speaker encounters trouble with their turn, their addressees escalate their use of non-verbal resources in order to help resolve the problem.

Here we extend this work by exploring the hypothesis that non-speaker gestures with the hands should be especially valuable at the points in conversation where problems with mutual-understanding are identified. As Bavelas has shown, listener gestures and facial expressions in particular can provide important content-specific information as well as serving generic interaction management functions such as who wants to speak next (Bavelas et al. 2000; Bavelas, Coates, and Johnson 2006)

Although self-repairs are the most common form of repair in conversation, occurring in more than a third of dialogue turns (Colman and Healey 2011) here we look here at clarification dialogues (also termed second position repair initiations) in which people explicitly address problems with a previous turn by directly requesting clarification. In ordinary dialogue clarification questions account for approximately 4% of dialogue turns (Purver, Ginzburg, and Healey 2003). Although self-repairs can also be prompted and addressed by the concurrent feedback from listeners, clarification dialogues represent a particularly explicit and critical form of feedback where the participants have, for one reason or another, decided that some point of understanding needs to be addressed before the conversation can proceed.

[1] This work looked at self-repairs within a speakers turn. These are also called *position one self-initiated self-repairs* in the Conversational Analytic literature (Schegloff, Jefferson, and Sacks 1977).

Hypotheses

There are two elements to the hypotheses explored here. First, that clarification sequences are situations in which the verbal component of the interaction is proving to be insufficient in some way and therefore additional non-verbal resources should be recruited to compensate.

Prediction 1: Gestures should be more frequent during clarification sequences than during non-clarification sequences.

Second, that if conversation is a collaborative endeavour, these should be situations in which (non-speaking) addressees should respond more to the attempt to help get the conversation back on track.

Prediction 2: Non-speaking addressees should gesture more than speakers during clarification sequences but not during non-clarification sequences.

In addition, we expect that if listeners are collaborating, or attempting to collaborate on resolving the content of the problematic material addressed by the clarification question and response then their responses in this context should involve an increased use of content specific gestures such as iconic and pantomime gestures (see below for definitions) and not just, for example, hand movements involved in generic feedback, word searches or attempts to bid for the conversational floor.

Prediction 3: Non-speaking addressees should produce more content-specific gestures during clarification sequences than during non-clarification turns.

To test these predictions we examine the patterns of speaker and hearer gesture during clarification sequences in a corpus of spontaneous story-telling interactions (Plant and Healey 2012).

Methods

The basic task, a variant of Bavelas' close-call story technique, involves pairs of people who are asked to recall a specific experience they've had and describe it to the other person. The experiences are prompted, randomly, from a set of six common experiences: a backache, a laugh, a massage, a stomachache, a toothache and a yawn. Participants alternate in describing these experiences to each other. Two additional experience prompts, a headache and a meal, were used as practice trials for all participants and were not analysed.

This task produces relatively free dialogue in which the person describing the experience tends to have most of the initiative but with some generic and specific feedback, questions and responses from their listener. Because the experiences all refer to something related to the body there tend to be more iconic and deictic gestures (see below) than is characteristic of other kinds of unscripted dialogue.

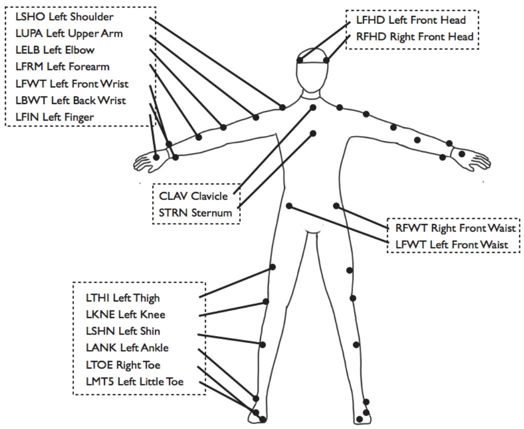

Figure 1: Position of reflective markers for motion-capture

Participants

12 pairs of participants took part, 12 females and 12 males with ages ranging from 18 to 60. The dyads consisted of a mixture of mixed sex and same sex pairs. 2 both-female, 3 both-male, and 5 mixed. Half of the dyads already knew each other and half were strangers prior to the study.

Materials and Apparatus

The interactions were recorded in a laboratory fitted with an optical based Vicon motion-capture system, consisting of 12 near infrared cameras. Participants wore a lycra top, trousers and headband with 41 reflective markers attached (see Figure 1 – note that the markers on the back are not shown). Near infrared cameras detect the optical markers at 60 frames per second, resulting in a highly accurate 3D representation of participants' movements over time. All interactions were also videoed with full body face on views of each participant (see Figure 2).

The prompts were provided by a set of cards placed on a small table next to where the participants stood. Each participant was given a stack of these cards, randomised before they began interacting, and were asked to take turns selecting one card at a time.

Procedure

Participants were given written instructions. They were asked to recall a specific instance of the prompted experience and explain it to their partner for no longer than 2 minutes. Emphasis was placed on describing how this experience felt at the time. The listeners were also encouraged to talk and ask questions at any time.

Motion Analysis

The optical motion capture data provides a quantitative index of hand movements. It is high resolution and detects even very fine-grained motions. In order to obtain an index of 'significant' motion we use a threshold of a change in any direction of the fastest moving hand that is more than one standard deviation from the mean movement for that person. This is coded frame-by-frame with hand movements (mm/frame) for each participant.

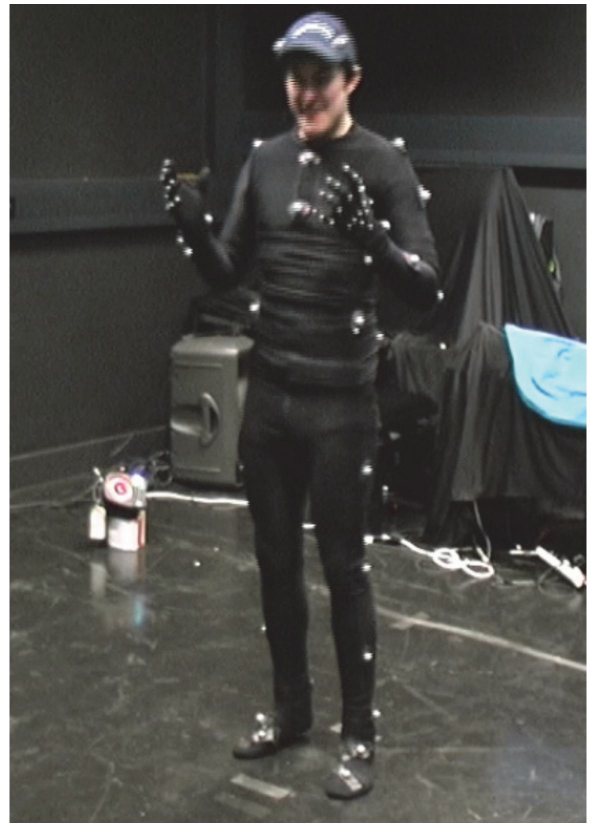

Figure 2: Still from video of one participant

Gesture Coding

The video and audio descriptions of each experience were imported into the ELAN coding tool and transcribed. Content-specific gestures were coded in two passes. On the first pass, any occurrences of physiographic gesture were coded without specifying their nature. On the second pass, content-specific gestures were separated into five types following (McNeill 2008; Rowbotham et al. 2012; Kendon 2004): Iconic, Metaphoric, Deictic, Pantomime and Abstract Descriptive following the definitions below. On each pass, only one camera view was coded at a time so coding for one participant's gestures could not influence the coding of the other's.

1. *Iconic* - a depiction of a concrete characteristic intrinsic to what is being described, such as size or shape. For example a gesture that makes a fist to represent roundness by making the hand round.

Figure 3: Illustration of Changes in Speaker and Non-Speaker Body Movements During a Clarification Sequence

2. *Metaphoric* - unlike iconic gestures metaphoric gestures depict abstract ideas rather than a concrete object or event.
3. *Deictic* - a gesture that points to the object of talk in order to locate it for the listener. Deictic gestures can be used to locate something that cannot be physically pointed to but can be referred within the speakers gesture space.
4. *Pantomime* - a re-enactment of behaviour that is intrinsic to the original action, like acting out the bodily reaction to stubbing a toe by hopping around in pain.
5. *Abstract Descriptive* - a specialised type of gestures that describe the characteristics of felt experiences such as pain, not including pointing to the location of the pain or miming the bodily reaction to the pain. For example, a gesture that represents the rhythmic quality of a throbbing pain by mimicking it in the rhythm of the hand movement, or perhaps the intensity intrinsic to the pain would be depicted by the tenseness in the fingers.

Feedback Coding

In order to contrast the production of content specific gestures with the provision of more generic non-speech feedback such as nods each instance of non-speech feedback was also coded in Elan according to three typical functions of addressee feedback (Allwood et al. 2007):

1. Contact and Perception (CP), indicating listener contact and perception of message
2. Comprehension (C), indicating listener comprehension or understanding of message
3. Attitudinal or Emotional (A/E), indicating an attitudinal or emotional response as simple as agreeing with the speaker (attitudinal), or showing shock to the speaker message (emotional), like motor mimicry.

Turn Coding

Dialogue was transcribed by utterance. In the analysis a person is coded as speaker for all occasions in which they are speaking including anything spoken in overlap with the other participant. Clarification sequences were additionally coded into two stages. The first stage was the question itself – what is asked when someone fails to fully comprehend something in another person's previous utterance (coded as CQ). The second stage was the response to the clarification question (coded as R). Responses to clarification questions are not always verbal – they can be in the form of a nod or a point toward the listener to confirm they are correct. Example clarification sequences are provided in Examples 1, 2 and 3.[2] An example exchange illustrating both speaker and hearer gestures is provided in Figure 3.

(1)

A: . . . and that movement really cracks your back

B: What's that? You do that and someone pulls? **CQ**

[2] In the examples '=' indicates a very short gap between utterances, numbers in parentheses '(0.1)' indicate pause lengths in 10ths of a second and '[ ]' on adjacent lines indicate overlapping text.

A: Yeah, you do that and someone just lifts you up from your shoulders there **R**

(2)

A: and all the tension goes up to the (0.1)

A: le (0.1)

A: upper back=

B: =so you are hurting the rest of [your body?] Yeah?= **CQ**

A: [Yeah]

A: =in the middle in the middle **R**

(3)

A: cos there is like, there something er (0.3)

A: it's not hurting though (0.2)

A: it's like=

B: sensitive? (0.1) **CQ**

A: yeah a little bit yeah (0.2) **R**

B: uh huh

Speaker Coding

In the analysis we distinguish between a) two task roles: 'story-teller' for the person relating their experience on a particular trial and 'story-recipient' for the person who is the audience for the story and b) two dialogue roles: 'speaker' for whoever is currently talking and 'non-speaker' or 'addressee' for a person who is not speaking. It is important to keep these distinct because this is a free dialogue and although the story-teller is typically also the speaker the story-recipient also produces verbal comments and feedback. On these occasions the 'story-recipient' becomes the speaker and the 'story-teller' becomes the addressee.[3]

Results

There are 4425 spoken turns transcribed in the corpus as a whole and 1349 content-specific gestures.

Gesture Types and Feedback

The combined results for both content-specific gestures and more generic forms of feedback across all turns show that people are active both when speaking and not speaking and global comparison of the frequency of non-speech signals that overlap with the construction of a turn shows no reliable difference due to dialogue roles ($\chi^2_{(2)}$= 0.78, p = 0.38).[4] This highlights the high level of collaboration between speakers and non-speakers during turn construction.

As Table 1 shows, although non-speakers are active they do not contribute the same types of non-speech signals as speakers. Speakers produce around 80% of the content specific gestures that overlap with a speakers own turn whereas non-speaking addressees produce around 80% of the non-verbal feedback that overlaps with a speaker's turn. This fits the general pattern reported in the literature that attentive non-speakers provide frequent generic 'backchannel' feedback to speakers to signal continued attention and interest and to support the speaker in the production of a turn (e.g.(Bavelas et al. 2000; Bavelas, Coates, and Johnson 2006)). Note that although speakers cannot sensibly provide concurrent feedback to themselves these results also suggest that they do still provide some non-speech feedback signals to the non-speaker while they are talking.

Table 1: Distribution of Feedback and Content Specific Gestures by Dialogue Role

Non-Speech Signals:	Content Specific	Feedback
Speaker:	2113 (79%)	550 (21%)
Non-Speaker:	553 (21%)	1975 (78%)

In the corpus a total of 194 turns are coded as clarification questions and 183 as responses. Of these the majority of clarification questions are initiated by the story recipient (166 = 85%) and the majority of responses are provided by the story teller (87%).

Prediction 1, that gestures should be more frequent overall during clarification dialogues, is not supported overall. Although slightly more clarification sequence turns overlap with content specific gestures 23% vs 20%) there is a trade-off between the dialogue roles; speakers gesture less frequently during clarification sequences than non-clarification sequences (dropping from 33% of turns overlapping with gestures to 26%) whereas non-speakers gesture more than double their content-specific gesture frequency (rising from 7% to 19%).

Prediction 2, that non-speakers should make more non-speech contributions than speakers during clarification sequences is also not directly supported (although see below). Counting both content-specific gestures and generic feedback there are 185 overlapping speaker gestures/feedback vs. 193 overlapping non-speaker gestures/feedback. However, this masks a significant shift by non-speakers from generic feedback to content specific gestures.

As Table 2 shows, the general pattern for speakers and non-speakers shown in Table 1 changes substantially during clarification sequences. Non-speaking addressees become substantially more likely to produce content-specific gestures in overlap with the speaker's turn (increasing from 21% to 49% of their gestures $\chi^2_{(1)}$= 0.734, $p < 0.001$) and speakers themselves become more likely to produce non-verbal feedback in overlap with their own turn (increasing from 21% to 31% of turns, $\chi^2_{(1)}$= 10.6, $p < 0.001$). This supports Prediction 3, showing that non-speakers make more use of content-specific gestures during clarification sequences.

Hand Movement Results

In addition to the gesture type analysis the motion capture data allows us to assess overall patterns of hand movement.

[3]Therefore our 'story recipient role corresponds to Gullberg's (1998) 'listener' whereas our 'addressee' corresponds to Furuyama's (2002) 'listener'.

[4]Throughout we report computed probabilities but adopt a criterion value of $p < 0.05$.

Table 2: Distribution of Feedback and Content Specific Gestures by Dialogue Role in Clarification Sequences

Non-Speech Signals:	Content Specific	Feedback
Speaker:	128 (69%)	57 (31%)
Non-Speaker:	94 (49%)	99 (51%)

To assess whether speakers normally move their hands more than listeners in this corpus, a Generalised Linear Mixed Model (GLMM) analysis of the average frequency of hand significant movements (as defined above) for each participant during all dialogue turns compared the hand movements with Speaking as a fixed factor (Yes vs. No) and Dyad as a random factor. This shows a main effect of Speaking (Yes or No) ($F_{(1,518)}$) = 96.1, p ¡ 0.001). Speakers in the dyads move their hands approximately 25% faster (5.42) than non-speakers (4.03) across all turns.

A second focused GLMM analysis of hand movements during clarification dialogues for each trial with Dyad as a Random Factor and Speaking (Speaker or Addressee), Clarification Stage (Question or Response) and Clarification Stage $\times$ Speaking interaction shows no simple main effects of Speaking ($F_{(1,143)}$) = 0.12, p = 0.73) or Clarification ($F_{(1,143)}$) = 0.58, p = 0.45) but a significant Clarification Stage $\times$ Speaking interaction ($F_{(1,143)}$) = 14.24, p $<$0.001). The interaction is illustrated in Figure 4. Overall these results shows that during the clarification dialogue as a whole the speed of speaker and non-speaker hand movements are not reliably different. However, they vary across the different stages of the clarification sequence. When the clarification question is asked the non-speakers move their hands more than speakers but during the response, speakers move their hands more than non-speakers. This provides qualified support for Prediction 2. When a clarification question is posed non-speaking addresses are moving their hands more than speakers but not when it is answered.

As noted above, clarification questions in this task are mostly, although not always, posed by the story-recipient (see above for terminology) which suggests that the story-teller does not immediately suspend their hand movements when a question is posed about something they've just said. This is different from the way they would normally reduce their hand movements if a normal change of speaker had taken place. During the response, again typically produced by the the story-teller, they continue to move their hands more quickly than the non-speaker. This suggests that ownership of the content being queried may have a significant influence on the organisation of hand movements at these points.

Discussion

The results reported here provide additional support for the claim that in free dialogue both speaking and non-speaking participants actively contribute to the production of each turn (Goodwin 1979; Clark 1996). Although they move their hands less than speakers, non-speakers provide frequent concurrent feedback to speakers and sometimes use non-speech signals to engage directly in helping the speaker to produce their turn (Tabensky 2001; Furuyama 2002). In general, speakers move their hands more and produce most of the content-specific gestures whereas non-speakers predominantly provide feedback to signal continued attention and understanding.

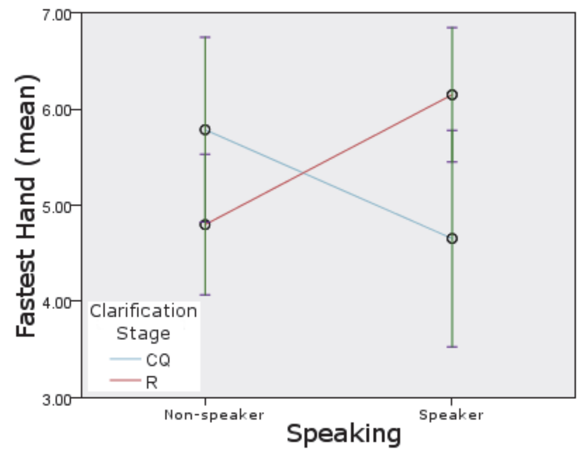

Figure 4: Estimated Marginal Means for The Interaction Between Clarification Stage and Speaking

This paper extends this literature by separating clarification sequences from the rest of the dialogues and showing that they involve a distinctive use of non-verbal resources in which both speakers and non-speakers change their behaviour. During clarification sequences speakers and non-speakers hand movements are no longer distinguishable in terms of speed. Non-speakers more than double the frequency of content-specific gestures they produce during clarification sequences and, although speakers still produce more content-specific gestures overall, speakers also tend to produce more frequent generic feedback. The dialogue roles that we conventionally distinguish on the basis of who is speaking thus become more blurred at these points as speaker and non-speaker collaboratively recruit additional, multi-modal, resources to deal with threats to mutual-understanding.

These results and the previous findings of (Healey et al. 2013) suggest a critical role for non-speaker contributions during repairs and clarifications and are consistent with the wider intuition that they represent the critical junctures in conversation. They support the claim that people are especially sensitive to these moments and, for non-speakers at least, that they will invest substantial additional effort in resolving them. They also suggest a potentially useful set of cues for automatically identifying critical points in natural interaction. It is unclear, of course, how much these findings may be specific to the particular task studied and the asymmetry in 'ownership' of content that the story-telling involves.

References

Allwood, J.; Cerrato, L.; Jokinen, K.; Navarretta, C.; and Paggio, P. 2007. The mumin coding scheme for the annotation of feedback, turn management and sequencing phenomena. *Language Resources and Evaluation* 273–287.

Bavelas, J. B.; Chovil, N.; Lawrie, D.; and Wade, L. 1992. Interactive gestures. *Discourse Processes* 15:469–489.

Bavelas, J. B.; Chovil, N.; Coates, L.; and Roe, L. 1995. Gestures specialized for dialogue. *Personality and social psychology bulletin* 21(4):394–405.

Bavelas, J. B.; Coates, L.; Johnson, T.; et al. 2000. Listeners as co-narrators. *Journal of personality and social psychology* 79(6):941–952.

Bavelas, J.; Coates, L.; and Johnson, T. 2006. Listener responses as a collaborative process: The role of gaze. *Journal of Communication* 52(3):566–580.

Clark, H. H., and Schaefer, E. A. 1989. Contributing to discourse. *Cognitive Science* 13:259–294.

Clark, H. H., and Wilkes-Gibbs, D. 1986. Referring as a collaborative process. *Cognition* 22:1–39.

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

Colman, M., and Healey, P. G. T. 2011. The distribution of repair in dialogue. In *Proceedings of the 33rd Annual Meeting of the Cognitive Science Society*, 1563–1568.

Emmorey, K., and Reilly, J. S., eds. 1995. *Language, gesture, and space*. Psychology Press.

Furuyama, N. 2002. Prolegomena of a theory of between-person coordination of speech and gesture. *International Journal of Human-Computer Studies* 57(4):347–374.

Goodwin, C. 1979. The interactive construction of a sentence in natural conversation. In Psathas, G., ed., *Everyday Language: Studies in Ethnomethodology*. New York: Irvington Publishers. 97–121.

Gullberg, M. 1998. *Gesture as a communication strategy in second language discourse: A study of learners of French and Swedish*, volume 35 of *Travaux de l'institut de linguistique de Lund*. Lund University.

Healey, P. G. T.; Lavelle, M.; Howes, C.; Battersby, S.; and McCabe, R. 2013. How listeners respond to speaker's troubles. In *Proceedings of the 35th Annual Conference of the Cognitive Science Society*.

Healey, P. G. T. 2008. Interactive misalignment: The role of repair in the development of group sub-languages. In Cooper, R., and Kempson, R., eds., *Language in Flux*. College Publications.

Holler, J., and Wilkin, K. 2011. Co-speech gesture mimicry in the process of collaborative referring during face-to-face dialogue. *Journal of Nonverbal Behavior* 35(2):133–153.

Kendon, A. 2004. *Gesture: Visible action as utterance*. Cambridge University Press.

McNeill, D. 2008. *Gesture and thought*. University of Chicago Press.

Mondada, L. 2014. Bodies in action: Multimodal analysis of walking and talking. *Language and Dialogue* 4(3):357–403.

Özyürek, A. 2002. Do speakers design their cospeech gestures for their addressees? the effects of addressee location on representational gestures. *Journal of Memory and Language* 46(4):688–704.

Plant, N., and Healey, P. G. 2012. The use of gesture to communicate about felt experiences. In *Proceedings of SemDial 2012 (SeineDial): The 16th Workshop on the Semantics and Pragmatics of Dialogue*, 80. Citeseer.

Purver, M.; Ginzburg, J.; and Healey, P. G. T. 2003. On the means for clarification in dialogue. In Smith, R., and van Kuppevelt, J., eds., *Current and New Directions in Discourse & Dialogue*. Kluwer Academic Publishers. 235–255.

Rowbotham, S.; Holler, J.; Lloyd, D.; and Wearden, A. 2012. How do we communicate about pain? a systematic analysis of the semantic contribution of co-speech gestures in pain-focused conversations. *Journal of Nonverbal Behavior* 36(1):1–21.

Schegloff, E.; Jefferson, G.; and Sacks, H. 1977. The preference for self-correction in the organization of repair in conversation. *Language* 53(2):361–382.

Schegloff, E. 1987. Some sources of misunderstanding in talk-in-interaction. *Linguistics* 25:201–218.

Schegloff, E. 1992. Repair after next turn: The last structurally provided defense of intersubjectivity in conversation. *American Journal of Sociology* 1295–1345.

Tabensky, A. 2001. Gesture and speech rephrasings in conversation. *Gesture* 1(2):213–236.

Yngve, V. H. 1970. On getting a word in edgewise. In *Papers from the 6th regional meeting of the Chicago Linguistic Society*, 567–578.