

# Using the Hands to Identify Who Does What to Whom: Gesture and Speech Go Hand-in-Hand

Wing Chee So,<sup>a,b</sup> Sotaro Kita,<sup>c</sup> Susan Goldin-Meadow<sup>a</sup>

<sup>a</sup>*Department of Psychology, University of Chicago*

<sup>b</sup>*Department of Psychology, National University of Singapore*

<sup>c</sup>*School of Psychology, University of Birmingham*

Received 15 January 2008; received in revised form 31 March 2008; accepted 1 May 2008

---

## Abstract

In order to produce a coherent narrative, speakers must identify the characters in the tale so that listeners can figure out who is doing what to whom. This paper explores whether speakers use gesture, as well as speech, for this purpose. English speakers were shown vignettes of two stories and asked to retell the stories to an experimenter. Their speech and gestures were transcribed and coded for referent identification. A gesture was considered to identify a referent if it was produced in the same location as the previous gesture for that referent. We found that speakers frequently used gesture location to identify referents. Interestingly, however, they used gesture most often to identify referents that were also uniquely specified in speech. Lexical specificity in referential expressions in speech thus appears to go hand-in-hand with specification in referential expressions in gesture.

*Keywords:* Gesture; Referential expression; Communication; Discourse

---

## 1. Introduction

When telling a narrative, speakers have to specify the characters in such a way that listeners can identify who is doing what to whom (Bosch, 1983; Garrod, 2001; Lyons, 1977). Speakers can use nouns, pronouns, or zero anaphora to indicate characters (e.g., Fox, 1987; Givon, 1984), but they must use these devices appropriately, that is, in line with pragmatic discourse principles. According to Grice (1975), speakers should make their speech clear but no more informative than is required by the context (Maxim of Quantity). For example,

---

Correspondence should be sent to Wing Chee So, National University of Singapore, BLK AS4 9 Arts Link.  
E-mail: psyswc@nus.edu.sg

pronouns, as opposed to nouns, are typically used to refer to previously mentioned referents. Using a noun rather than a pronoun could create comprehension difficulties for a listener who might (reasonably) assume that the noun was used in order to signal a new referent. However, at times, referents are left under-specified. For example, in a story about two male protagonists, saying “*he pulled him into the water*” or even “*the man pulled the other man into the water*” leaves it ambiguous as to who is pulling whom. In contrast, in a story about a male protagonist and a female protagonist, saying “*he pulled her into the water*” or “*the man pulled the woman into the water*” leaves no ambiguity.

The question we address in this study is whether speakers use gesture to help them specify referents when they fail to do so in speech. Take the ambiguous case with two male protagonists. Imagine that the speaker gestures while first introducing the characters: “*There’s a guy with a hat* (points to his right) *and a guy with a scarf* (points to his left).” The speaker can then use these gestures to disambiguate his next utterance: “*And he* (points to his right) *pulled him* (points to his left) *into the water.*”<sup>1</sup> The speaker’s gestures (but not his words) make it clear that the man with the hat is doing the pulling, and not the reverse.

All gestures are produced in space, but the space must be used consistently over discourse in order for gesture to function effectively to identify a referent (Gullberg, 1998, 2003, 2006). So, Coppola, Licciardello, and Goldin-Meadow (2005) explored speakers’ ability to use gesture to identify referents under two conditions: Native English speakers, all naïve to sign language, were first asked to describe a story in speech (they were free to gesture but were not told to gesture) and were then asked to describe the same story using gesture without any speech. So et al. found that the storytellers used the spatial locations of the gestures they produced to identify referents in both conditions, although gestural location was used less consistently in their co-speech gestures than in the gestures they produced without speech.

In the present study, we explore the circumstances under which speakers use the spatial locations of co-speech gesture to identify referents. Specifically, we ask whether speakers use gestures to identify referents more often when those referents are *not* uniquely specified in speech. We thus investigate how speakers semantically coordinate speech and gesture to disambiguate information that is crucial for discourse processing.

Previous work led us to consider two competing predictions. First, on the assumption that speakers gesture for the benefit of their listeners (e.g., Cohen, 1977; Kendon, 1983), de Ruiter (2006) has suggested that speakers gesture to maximize information. Thus, when information crucial to communication is not conveyed in speech, gesture ought to step in and convey the information instead (the cross-modal compensation hypothesis). According to this hypothesis, speakers should use gesture to identify a referent particularly when speech fails to uniquely specify the referent. If, for example, speakers specify referents uniquely in speech *less often* when telling a story about two same-gender protagonists than two different-gender protagonists, they ought to identify referents in gesture *more often* in the same-gender story than in the different-gender story—in other words, they should use gesture to compensate for their under-specification in speech.

Alternatively, although listeners may profit from the information in a speaker’s gestures (Goldin-Meadow & Sandhofer, 1999), speakers could be gesturing for their own cognitive

benefit. If so, their gestures may reflect how they organize their thinking for the purposes of speaking (Slobin, 1987, 1996). In fact, Kita and Özyürek (2003) have suggested that gesture originates from the interface between speaking and thinking, where spatio-motoric imagery is packaged into units that are suitable for speaking (the interface hypothesis). Under this view, when speech fails to uniquely specify a referent, gesture should also fail to identify the referent. We would therefore expect speakers to identify referents in gesture *less often* in a same-gender story than in a different-genders story, paralleling the pattern found in speech.

To test these competing views, we manipulated lexical specificity in referential expressions in speech by asking speakers to describe two stories, the Man-Woman (M-W) story involving protagonists of different genders, and the Man-Man (M-M) story involving protagonists of the same gender. We first established that speakers are indeed less likely to uniquely specify referents in speech in the M-M story than in the M-W story. We then explored whether and how the differences in referent specification in speech influenced referent identification in gesture. In particular, we examined whether gesture compensates for, or parallels, under-specification of referents in speech.

## 2. Method

### 2.1. Participants and procedure

Nine English-speaking undergraduate students from the University of Chicago, naïve to sign language,<sup>2</sup> were recruited through postings throughout the campus and paid for their participation. We used data from the six participants in the So et al. (2005) study who produced co-speech gestures when telling at least one of their stories, and collected data from three new participants.

The participants were individually shown 81 videotaped vignettes, each lasting 1–6 s, culled from eight silent stories (see So et al., 2005). Participants were first shown all of the vignettes in a given story without pauses so that they could get a sense of the plot. The vignettes in the story were then repeated one at a time and, after each vignette, participants were asked to describe the scene to an experimenter who also watched the vignettes; the participants thus saw each vignette twice. The entire session was videotaped.

We used data from two of the eight stories. The M-W story (also used in So et al., 2005) contained 11 vignettes displaying a variety of motion events (e.g., man gives woman a basket; woman walks upstairs; man kisses woman; man falls).<sup>3</sup> A second story, the M-M story (not used in So et al., 2005) contained 10 vignettes also displaying motion events (e.g., the first man drops a rock on the foot of the second man; the second man removes noose from the neck of the first man; the first man throws the second man into water).

The study was a within-subject design, with story (M-W vs. M-M) as the single factor. Data containing proportions were subjected to an arcsine transformation before statistical analysis.

## 2.2. Speech coding

We selected vignettes in which both protagonists appeared (see Appendix A) and analyzed all of the speech produced to describe the two protagonists. We classified all references to the protagonists as containing either a pronoun (e.g., he, she) or a noun (e.g., man, woman, goofy-looking man), and determined whether the pronoun or noun uniquely specified its referent. For example, in the sentence, “*The guy hands the girl the basket. He hands her the basket*” (M-W story), the speaker used two distinct pronouns that uniquely referred back to their referents (“he” uniquely referred back to the guy and “her” referred back to the girl). In contrast, in the sentence, “*The two men are facing each other and, like, the one with the nose drops a big rock on the other man’s, accidentally drops a big rock on the other man’s foot, and he grabbed his foot*” (M-M story), the pronouns did not uniquely specify the referents. The speaker used “he” to refer to one man and “his” to refer to the other, and did not make it clear who grabbed whose foot. Note that referential under-specification can, and did, happen with nouns as well as pronouns, e.g., “*The man grabbed the other man’s foot*” (M-M story).

## 2.3. Gesture coding

Gesture form was described in terms of the parameters used to describe sign languages (Stokoe, 1960)—shape and placement of the hand, trajectory of the motion. Change in any one of these parameters marked the end of one gesture and the beginning of another. Each gesture was assigned a semantic meaning indicating which protagonist was represented in the gesture.

We used criteria developed by Senghas and Coppola (2001) to code use of space (see also Padden, 1988). To determine whether a gesture was used to identify a referent, we assessed the spatial location of each gesture produced for a particular character (a gesture was classified as referring to a character if it co-occurred with speech referring to that character) in relation to the spatial location of previously produced gestures for that character. A gesture was considered to identify the referent if it was produced in the same location (left, right, center, top or bottom relative to the location of a speaker) as the previous gesture for that referent. The proportion of gestures that identified referents was calculated for each speaker, and as was the number of referent-identifying gestures that the speaker produced, divided by the total number of gestures the speaker produced for the two protagonists in each story.

Reliability was assessed by having a second coder transcribe a subset of both stories. Agreement between coders was 89% ( $n = 97$ ) for identifying gestures and describing their form, 95% ( $n = 87$ ) for assigning semantic meaning to gestures, and 98% ( $n = 87$ ) for determining whether a gesture for a referent was produced in the same location as the previously produced gesture for that referent (i.e., whether it identified a referent).

### 3. Results

We asked whether referent identification in gesture went hand-in-hand with lexical specificity in referential expressions in speech. Thus, we first analyzed lexical specificity in referential expressions in speech, followed by referent identification in gesture.

#### 3.1. Lexical specificity in referential expressions in speech

The mean number of times participants referred to the protagonists in speech was 19.0 ( $SD = 4.5$ ) in the M-M story and 14.9 ( $SD = 5.8$ ) in the M-W story. Overall, the participants uniquely specified the referents in the two stories 81% ( $SD = 8%$ ) of the time. However, as expected, participants were significantly less likely to uniquely specify referents in the same-gender M-M story ( $M = 70%$ ,  $SD = 11%$ ) than in the different-genders M-W story ( $M = 99%$ ,  $SD = 1%$ ),  $t(8) = 5.56$ ,  $p < .0001$ . All nine of the participants displayed this pattern.

Speakers used nouns (as opposed to pronouns) to refer to the protagonists more often in the M-M story (60%,  $SD = 15%$ ) than in the M-W story (19%,  $SD = 17%$ ),  $t(8) = 6.57$ ,  $p < .0001$ . However, those nouns were used less precisely in the M-M than the M-W story: 80% ( $SD = 18%$ ) of the nouns that were used in the M-M story uniquely specified their referents, compared to 100% ( $SD = 0$ ) of the nouns in the M-W story,  $t(6) = 5.04$ ,  $p < .003$ . Seven participants displayed this pattern; the remaining two produced no nouns in the M-W story. Participants also used pronouns less precisely in the M-M story than in the M-W story: 53% ( $SD = 15%$ ) of pronouns uniquely specified their referents in the M-M story, compared to 99% ( $SD = 2%$ ) of pronouns in the M-W story,  $t(8) = 6.23$ ,  $p < .0001$ . Eight of nine participants displayed this pattern. Overall, the speakers in our study successfully used speech to uniquely specify the characters in their stories, but they did so less often in the M-M story than in the M-W story.

It is possible, however, that our codes were not sensitive to all of the ways in which the speakers had used speech to identify referents. We therefore administered a speech comprehension task to a naïve group of listeners ( $n = 20$ ) to verify the findings. The listeners did not watch the stories and were presented only with the audio (not the video) portion of the descriptions generated by the nine speakers. After each story, the listeners were given a series of test sentences containing blanks for the words that referred to the two protagonists (e.g., \_\_\_ gives \_\_\_ a basket) and were asked to fill in the blanks. The maximum comprehension score possible was 5 for each story (participants had to be correct on both blanks to receive credit for a sentence). Although listeners were quite successful at identifying protagonists overall, they identified significantly fewer protagonists in the M-M story ( $M = 4.20$ ,  $SD = 0.23$ ) than the M-W story ( $M = 4.56$ ,  $SD = 0.27$ ),  $t(8) = 2.62$ ,  $p < .034$ . This pattern was found for the stories produced by seven of the nine speakers (listener scores were equal for the two stories for one speaker, but were higher for the M-M story than for the M-W story for the other speaker).

These findings thus verify our initial results: The storytellers did, indeed, use speech to uniquely specify referents less often in the M-M story than in the M-W story. We next asked

whether the speakers used gesture to compensate for the under-specification in speech in the M-M story.

### 3.2. Referent identification in gesture

On average, speakers produced 14.10 ( $SD = 7.49$ ) gestures when referring to protagonists in the M-M story and 9.56 ( $SD = 6.97$ ) in the M-W story (one participant produced no gestures at all in the M-W story). Speakers produced gestures along with 43% ( $SD = 22%$ ) of the words they produced to refer to the two protagonists in the M-M story and 44% ( $SD = 31%$ ) in the M-W story.

Moreover, speakers used the spatial location of their gestures systematically to identify referents: They produced gestures in locations previously identified with a particular character 35% ( $SD = 26%$ , range from 5% to 52%) of the time. The crucial question, however, is whether speakers used gesture to identify referents more often in the M-M story (where speech under-specified referents) than in the M-W story. They did not. In fact, speakers identified referents in gesture reliably *less often* when telling the M-M story (27%,  $SD = 14%$ ) than when telling the M-W story (62%,  $SD = 30%$ ),  $t(7) = 4.23$ ,  $p < .009$ ; seven of the nine speakers displayed this pattern.

Why were speakers less likely to use gestures to identify referents in the M-M story than in the M-W story? One possibility is that gesture goes hand-in-hand with speech, identifying referents only when speech does. To explore this hypothesis, we looked at whether participants used gesture to identify a particular referent when that referent was, and was not, uniquely specified in speech. Across both stories, speakers were reliably more likely to identify a referent in gesture when they also uniquely identified that referent in speech ( $M = 45%$ ,  $SD = 26%$ ) than when they did not uniquely identify the referent in speech ( $M = 9%$ ,  $SD = 23%$ ),  $t(7) = 3.05$ ,  $p < .019$ ; seven of the nine speakers displayed this tendency. Moreover, this pattern was found in both stories: in the M-W story, 62% ( $SD = 31%$ ) of referents that were uniquely identified in speech were identified in gesture, compared to none of the referents that were *not* uniquely identified in speech; comparable percentages for the M-M story were 34% ( $SD = 30%$ ) and 9% (22%).

Interestingly, gesture was rarely used to compensate for the absence of lexical specificity in pronouns or nouns. As shown in Fig. 1, only 13% ( $SD = 36%$ ) of the pronouns that did *not* uniquely specify a particular referent were accompanied by a gesture that identified the referent, compared to 55% ( $SD = 20%$ ) of the pronouns that did specify a particular referent,  $t(6) = 1.94$ ,  $p < .05$ . The same pattern was found for nouns: only 4% ( $SD = 15%$ ) of the nouns that did *not* uniquely specify a particular referent were accompanied by a gesture that identified the referent, compared to 21% ( $SD = 10%$ ) of the nouns that did specify a particular referent,  $t(6) = 2.23$ ,  $p < .018$ . In other words, very rarely did gesture specify the identity of a referent if that referent was not also uniquely specified in speech.

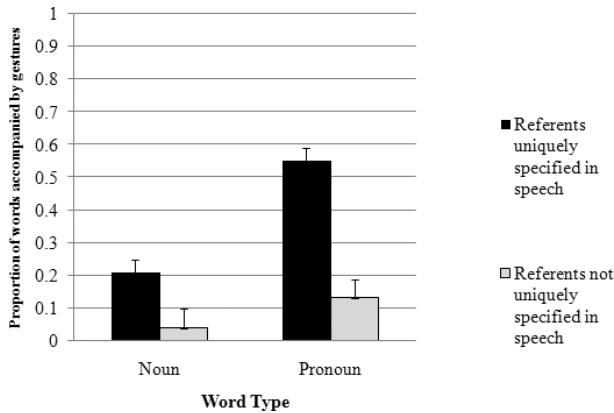


Fig. 1. Proportions of words that were accompanied by referent-identifying gestures. The black bars represent words that uniquely specified the referent; the white bars represent words that did not. Nouns are displayed in the left bars, pronouns in the right bars. Both nouns and pronouns were more likely to be accompanied by gestures when they uniquely specified the referents than when they did not.

#### 4. Discussion

Our study asked whether speakers use gesture to identify referents that are not specified in speech, or whether gesture is so tightly linked to speech that it identifies referents only when speech does. We found that 35% of the gestures speakers produced were placed in locations previously associated with a character. Thus, speakers did use gesture to specify the identity of a referent. However, they did not use gesture to specify the identity of a referent unless that referent was also uniquely specified in speech. In other words, they did not use gesture to compensate for the under-specification found in their speech. Rather, they used gesture in parallel with their speech.

Although speakers could have used gesture to specify a referent's identity when that referent was not uniquely specified in speech, they rarely did. The question is *why not?* We suggest that speakers did not use gesture to compensate for speech because gesture and speech are part of a single, integrated system (McNeill, 1992). Previous work has shown that gestures tend to convey information that is semantically coordinated with the information conveyed in the speech those gestures accompany (the interface hypothesis, Kita, 2000; Kita & Özyürek, 2003; Özyürek & Kita, 1999; Özyürek, Kita, Allen, Furman, & Brown, 2005). For example, English has a word to convey an arced trajectory (“swing”) but Turkish and Japanese do not. When describing an arced trajectory, English speakers typically use the word “swing” and produce gestures with an arced trajectory. Japanese and Turkish speakers not only do not have a word like “swing” at their disposal to describe the event but, interestingly, they do not produce gestures with an arced trajectory to compensate for the lexical gap in their vocabularies. They produce speech referring to a change of location without specifying the trajectory (e.g., a word like “go” in English) and a parallel gesture with a straight trajectory despite the fact that the path they are describing is arced (Kita & Özyürek, 2003). Thus, speakers tend to produce gestures that parallel the lexical

patterns in their speech, just as, in our study, speakers produced gestures that parallel the lexical specification patterns in their speech.

Our findings also bear on the level at which gesture and speech must be integrated. The fact that gesture is used to track the identity of a referent across a piece of discourse suggests that gesture is linked to speech not only at the word or sentence level but also at the discourse level (Levy & McNeill, 1992). If speakers were to coordinate their gestures with speech only at the word or sentence level (and not at the discourse level), we would not find, as we do, gestures that maintain the location of a referent from one sentence to the next.

If gesture and speech do go hand-in-hand, why is it then that speakers use gesture to identify a referent only 35% of the time? In other words, why don't speakers use gesture to identify a referent every time they specify that referent in speech? Gesture is tightly integrated with the speech it accompanies to form a communication system and, as a result, it is called upon to serve multiple functions. In addition to identifying referents, gesture is used to convey propositional information (Goldin-Meadow, 2003), coordinate social interaction (Bavelas, Chovil, Lawrie, & Wade, 1992; Haveland, 2000), and break discourse into chunks (Kendon, 1972; McNeill, 2000b). Thus, using gesture to identify referents in any particular discourse unit is likely to have to compete with a variety of other functions that gesture serves.

A priori we might have guessed that speakers would use gesture to convey information that is not found in their speech (the cross-modal compensation hypothesis, de Ruyter, 2006). But the speakers in our study did *not* use gesture to compensate for the referents they failed to specify in speech. These findings are consistent with the hypothesis that speakers do not gesture *solely* for interactional or communicative purposes. Work by Gullberg (2006) on gesture's role in referent tracking supports this hypothesis (see also Iverson & Goldin-Meadow, 1998; Rimé, 1982). Gullberg manipulated whether speakers could see their listeners, and found that speakers continued to use gesture to identify referents even when those gestures were not visible to the listener. In other words, the referent-identifying information that speakers included in their gestures was not exclusively for the listener's benefit, else it would have disappeared when the listener could no longer see those gestures.

Evidence that gesture and speech go hand-in-hand has, up until this point, come mainly from cross-linguistic analyses (but see Kita et al., 2007); speakers of different languages have been found to gesture differently and in accord with differences between their languages. Our data add support to the hypothesis by looking not only within-culture but also within-speaker. We found that when English speakers fail to uniquely identify a referent in speech, they also fail to identify the referent in gesture. Once again, we find that gesture is used redundantly with speech.

But gesture can, at times, convey information that is *not* found in the speech it accompanies (Goldin-Meadow, 2003). In these cases, the speaker is typically either in an unstable cognitive state, on the verge of acquiring the task (Goldin-Meadow, Alibali, & Church, 1993), or interacting with someone who is in an unstable state (Goldin-Meadow & Singer, 2003). The fact that speakers in our study rarely identified a referent in gesture when they failed to identify the referent in speech suggests that they had a stable understanding of



lexical specification. We might expect that children on the verge of mastering lexical specification in speech would use gesture differently, frequently identifying referents in gesture even though they are not yet able to identify them in speech.

To summarize, we found that speakers can use gesture to identify referents across a stretch of discourse. However, they rarely use gesture to identify a referent unless that referent is also uniquely specified in speech. Thus, referent identification in gesture appears to go hand-in-hand with lexical specificity in speech.

## Notes

1. In this example, it is likely that the speaker would stress the pronouns “he” and “him.” Indeed, it is often the case that speakers stress pronouns that are accompanied by gesture (Marslen-Wilson, Levy, & Tyler, 1982).
2. Conventional sign languages such as American Sign Language use space grammatically as part of the pronominal system (Padden, 1988). We therefore included in our study only participants who had no knowledge of sign language to avoid the possibility that the way they used space in their co-speech gestures might have been influenced by sign.
3. The analyses in the present study differ from those in So, Coppola, Licciardello, & Goldin-Meadow (2005) in that here we looked only at references to animate characters. In So et al., references to inanimate objects were also included in the analyses.

## Acknowledgments

This research was supported by R01 DC00491 from National Institute on Deafness and other Communication Disorders to Susan Goldin-Meadow, and Grant No. SBE0541957 from National Science Foundation to support the Spatial Intelligence and Learning Center at the University of Chicago. We thank Jessie Moth and Tsui Hoi Yan for their help in establishing reliability in speech and gesture coding.

## References

- Bavelas, J. B., Chovil, N., Lawrie, D. A., & Wade, A. (1992). Interactive gestures. *Discourse Processes*, 15, 469–489.
- Bosch, P. (1983). *Agreement and anaphora: A study of the roles of pronouns in discourse and syntax*. London: Academic Press.
- Cohen, A. A. (1977). The communicative functions of hand illustrators. *Journal of Communication*, 27, 54–63.
- Fox, B. (1987). *Discourse structure and anaphora*. Cambridge, England: Cambridge University Press.
- Garrod, S. (2001). Anaphora resolution. In N. J. Smelser & P. B. Baltes (Eds.), *International encyclopedia of the social and behavior sciences* (pp. 490–494). Amsterdam: Elsevier.
- Givon, T. (1984). Universals of discourse structure and second language acquisition. In W. E. Rutherford (Ed.), *Language universals and second language acquisition* (pp. 109–136). Amsterdam: Benjamins.

- Goldin-Meadow, S. (2003). *Hearing gesture: How our hands help us think*. Cambridge, MA: Harvard University Press.
- Goldin-Meadow, S., Alibali, M. W., & Church, R. B. (1993). Transitions in concept acquisition: Using the hand to read the mind. *Psychological Review*, *100*, 279–297.
- Goldin-Meadow, S., & Sandhofer, C. M. (1999). Gesture conveys substantive information about a child's thoughts to ordinary listeners. *Developmental Science*, *2*, 67–74.
- Goldin-Meadow, S., & Singer, M. A. (2003). From children's hands to adults' ears: Gesture's role in teaching and learning. *Developmental Psychology*, *39* (3), 509–520.
- Grice, H.P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and semantics*, Volume 3 (pp. 41–58). New York: Academic Press.
- Gullberg, M. (1998). *Gesture as a communication strategy in second language discourse: A study of learners of French and Swedish*. Lund: Lund University Press.
- Gullberg, M. (2003). Gestures, referents, and anaphoric linkage in learner varieties. In C. Dimroth & M. Starren (Eds.), *Information structure, linguistic structure and dynamics of language acquisition* (pp. 311–328). Amsterdam: Benjamin.
- Gullberg, M. (2006). Handling discourse: gestures, reference tracking, and communication strategies in early L2. *Language Learning*, *56* (1), 156–196.
- Haviland, J. (2000). Pointing, gesture spaces, and mental maps. In D. McNeill (Ed.), *Language and gesture* (pp. 13–46). New York: Cambridge University Press.
- Iverson, J. M., & Goldin-Meadow, S. (1998). Why people gesture as they speak. *Nature*, *396*, 228.
- Kendon, A. (1972). Some relationships between body motion and speech: An analysis of an example. In A. Siegman & B. Pope (Eds.), *Studies in dyadic communication* (pp. 177–210). Elmsford, NY: Pergamon Press.
- Kendon, A. (1983). Gesture and speech: How they interact. In J. M. Weimann & R. P. Harrison (Ed.), *Nonverbal interaction*. Beverly Hills, CA: Sage.
- Kita, S. (2000). How representational gestures help speaking. In D. McNeill (Ed.), *Language and gesture* (pp. 162–185). Cambridge: Cambridge University Press.
- Kita, S., & Özyürek, A. (2003). What does cross-linguistic variation in semantic coordination of speech and gesture reveal? Evidence for and interface representation of spatial thinking and speaking. *Journal of Memory and Language*, *48*, 16–32.
- Kita, S., Özyürek, A., Allen, S., Brown, A., Furman, R., & Ishizuka, T. (2007). Relations between syntactic encoding and co-speech gestures: Implications for a model of speech and gesture production. *Language and Cognitive Processes*, *22* (8), 1212–1236.
- Levy, E. T., & McNeill, D. (1992). Speech, gesture, and discourse. *Discourse Processes*, *15*, 277–301.
- Lyons, J. (1977). *Semantics 2*. Cambridge, England: Cambridge University Press.
- Marslen-Wilson, W., Levy, E., & Tyler, L. K. (1982). Producing interpretable discourse: The establishment and maintenance of reference. In R. Jarvella (Ed.), *Speech, Place, and Action* (pp. 339–378). Hoboken, NJ: John Wiley & Sons Ltd.
- McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago: University of Chicago Press.
- McNeill, D. (2000b). Catchments and contexts: Non-modular factors in speech and gesture production. In D. McNeill (Ed.), *Language and gesture* (pp. 312–328). NY: Cambridge University Press.
- Özyürek, A., & Kita, S. (1999). Expressing manner and path in English Turkish: Differences in speech, gesture, and conceptualization. In M. Hahn & S. C. Stonnes (Eds.), *Proceedings of the twenty-first annual meeting of the cognitive science society* (pp. 507–512). Hillsdale, NJ: Erlbaum.
- Özyürek, A., Kita, S., Allen, S., Furman, R., & Brown, A. (2005). How does linguistic framing of events influence co-speech gestures? *Gesture*, *5*, 219–240.
- Padden, C. (1988). *Interaction of morphology and syntax in American Sign Language*. New York: Garland.
- Rimé, B. (1982). The elimination of visible behaviour from social interactions: Effects on verbal, nonverbal and interpersonal variables. *European Journal of Social Psychology*, *12*, 113–129.
- de Ruiter, J.P. (2006). Can gesticulation help aphasic people speak, or rather, communicate? *Advances in Speech-Language Pathology*, *8* (2), 124–127.

- Senghas, A., & Coppola, M. (2001). Children creating language: How Nicaraguan Sign Language acquired a spatial grammar. *Psychological Science*, 12, 323–328.
- Slobin, D. (1987). Thinking for speaking. In J. Aske, N. Beery, L. Michaelis & H. Filip (Eds.), *Proceedings of the thirteenth annual meeting of the Berkeley linguistic society* (pp. 435–445). Berkeley, CA: Berkeley Linguistic Society.
- Slobin, D. I. (1996). From “thought and language” to “thinking to speaking.” In J. J. Gumperz & S. C. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 70–96). Cambridge, England: Cambridge University Press.
- So, W. C., Coppola, M., Licciardello, V., & Goldin-Meadow, S. (2005). The seeds of spatial grammar in the manual modality. *Cognitive Science*, 29, 1029–1043.
- Stokoe, W. C. (1960). Sign language structure: an outline of the visual communication systems of the American Deaf. *Studies in Linguistics, Occasional Papers* 8.

## Appendix A

---

### Man-Woman story

- |                |  |
|----------------|--|
| 1              | cat wags tail while sitting on windowsill        |
| 2              | man gets out of car while woman sits idly by     |
| 3 <sup>a</sup> | man holding a basket leads woman down a corridor |
| 4 <sup>a</sup> | man doffs hat towards woman                      |
| 5 <sup>a</sup> | man gives woman a basket                         |
| 6 <sup>a</sup> | man grabs woman’s hand                           |
| 7 <sup>a</sup> | man kisses woman’s hand                          |
| 8 <sup>a</sup> | woman walks up steps while man watches her       |
| 9              | woman opens door                                 |
| 10             | cat knocks pot off windowsill                    |
| 11             | falling pot strikes man on head and man falls    |

### Man-Man story

- |                 |   |
|-----------------|---|
| 1               | man holding a suitcase descends a flight of steps   |
| 2               | man opens suitcase  |
| 3               | man throws noose around his neck and then opens it  |
| 4 <sup>a</sup>  | man continues to hold noose around neck while he watches a second man with a cane descend the steps |
| 5 <sup>a</sup>  | first man drops a rock on the foot of the second man  |
| 6 <sup>a</sup>  | second man removes noose from first man   |
| 7 <sup>a</sup>  | first man throws noose back around his neck as well as around the second man                        |
| 8 <sup>a</sup>  | first man falls from under noose while second man looks off in the other direction                  |
| 9 <sup>a</sup>  | first man throws rock into a nearby body of water   |
| 10 <sup>a</sup> | second man pulls first man into water with him  |
- 

<sup>a</sup>Vignettes included in the analysis.