Production of speech-accompanying gesture

Sotaro Kita

University of Birmingham

Introduction

People spontaneously produce gestures when they speak. Gesture production and speech production are tightly linked processes. Speech-accompanying gesture is a cultural universal (Kita, 2009). Whenever there is speaking, there is gesture. Infants in the one-word stage already combine speech and gesture in a systematic way (Capirci, Iverson, Pizzuto, & Volterra, 1996; Iverson & Goldin-Meadow, 2005). Gesturing persists in situations where gestures are not communicatively useful, for example, when talking on the phone (J. Bavelas, Gerwing, Sutton, & Prevost, 2008; Cohen, 1977). Congenitally blind children spontaneously produce gestures (Iverson & Goldin-Meadow, 2001), indicating gesture is resilient against poverty of input.

Speech-accompanying gestures come in different types. The most influential classification system by McNeill (1992) distinguishes iconic (metaphoric) gestures, deictic gesture, beat gesture and emblem gestures. Iconic gestures can depict action, events and shapes in an analogue and iconic way (e.g., a hand swinging as if to throw a ball can represent throwing, a flat hand moving downward can represent a flat object falling, or a hand can represent a shape by tracing the outline). Such gestural depiction can also represent abstract contents by spatializing them (e.g., the flow of time can be represented by a hand moving across). Iconic gestures with abstract contents are sometimes given a different label, metaphoric gesture (Cienki & Müller, 2008; McNeill, 1992). Deictic (pointing) gestures indicate the referent by means of spatiotemporal contiguity (Kita, 2003). Beat gestures are small bi-directional movements that are often performed in the lower periphery of gesture space (e.g., near the lap) as if to beat the rhythm. The form of beat gestures remains more or less the same, regardless of the content of the concurrent speech. One of the proposed functions is to mark shifts in discourse structure (McNeill, 1992). Emblem gestures have a conventionalised and often arbitrary form-meaning relationship (e.g., the OK sign with a ring created by the thumb and the index finger) (Kendon, 1992; Morris, Collett, Marsh, & O'Shaughnessy, 1979). In the remainder of the chapter, the focus will be on iconic and deictic gestures (i.e., "representational gestures") because the bulk of psycholinguistic work on production has been on these two types of gestures (but see Krahmer & Swerts, 2007 for work on beat gestures).

A model of speech and gesture production

General architecture

Many of the empirical findings about speech-accompanying gestures can be explained by a model in which speech production and gesture production are regarded as separate but highly interactive processes such as in Figure 1 (Kita & Özyürek, 2003). This model is based on Levelt's (1989) model of speech production. The goal of this chapter is to provide an overview of the literature on speech-accompanying gestures, using the model as a means to organise information.

In Figure 1, the rectangles represent information processing components and arrows represent how the output of one processing component is passed on to another component. The ovals represent information storage and dotted lines represent an access route to information storage.

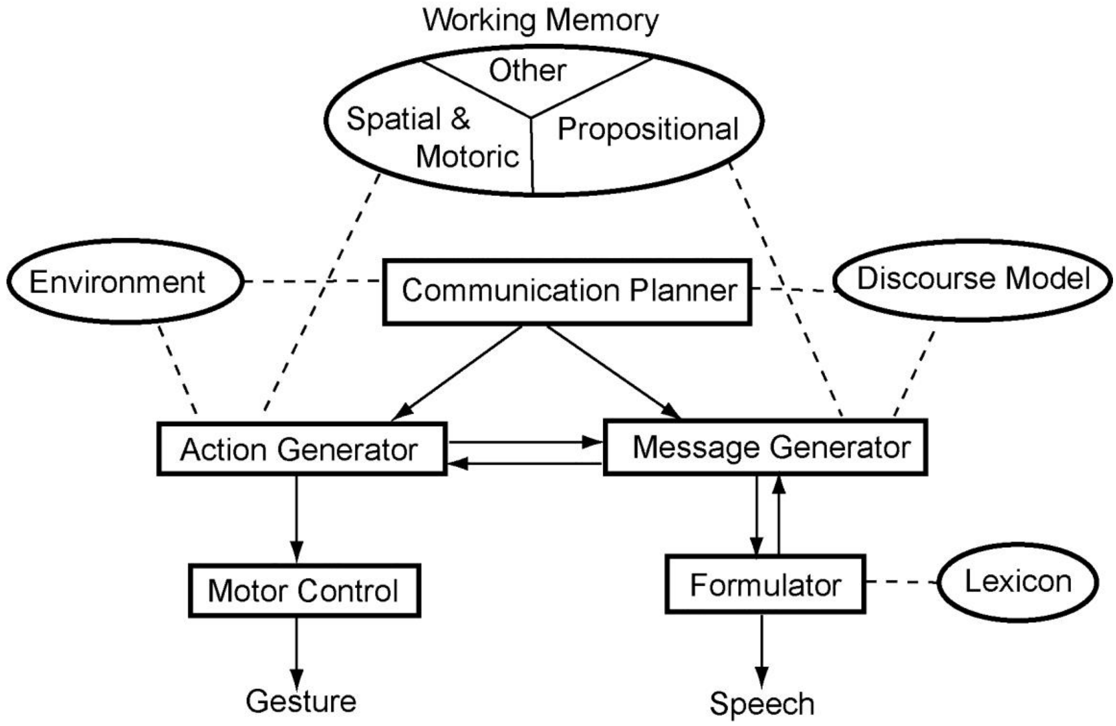

Figure 1. A model of speech and gesture production (Kita & Özyürek, 2003) (permission pending).

As in Levelt (1989), two distinct planning levels for speech production are distinguished. The first concerns planning at the conceptual level ("Conceptualizer"), which determines what message should be verbally encoded. The content of the message is determined on the basis of what is communicatively needed and appropriate, based on information about the discourse context (Discourse Model) and on the relevant propositional information activated in the working memory. The second concerns planning of linguistic formulation ("Formulator"), which linguistically encodes the message. That is, it specifies the words to be used, the syntactic relationship among the words and the phonological contents of the words.

Levelt's Conceptualizer is divided into the Communication Planner and the Message Generator. The Communication Planner corresponds to "macroplanning" in

Levelt's model. This process determines roughly what contents need to be expressed (i.e., communicative intention) in what order. In addition, the Communication Planner determines which modalities of expression (speech, gesture) should be used for communication (see de Ruiter, 2000 for a related idea that Conceptualizer determines which modalities of expression should be used), taking into account the extent to which the Environment is suitable for gestural communication (e.g., whether or not the addressee can see the speaker's gesture). Thus, the Communication Planner is not dedicated to speech production, but plans multi-modal communication as a whole. The Message Generator corresponds to "microplanning" in Levelt's model. This process determines precisely what information needs to be verbally encoded (i.e., preverbal message).

Gesture production follows similar steps to speech production. At the gross level, two distinct levels of planning are distinguished. The Communication Planner and the Action Generator carry out the conceptual planning for gesture production and the Motor Control execute the conceptual-level plans. The Communication Planner determines roughly what contents need to be expressed in the gesture modality. The Action Generator determines precisely what information is gesturally encoded. The Action Generator is a general-purpose process that plans actions in real and imagined environments.

In the following sections, we will discuss interaction between the components in the model. We will start with the description of how the Communication Planner and the Action Generator work. Then, we will discuss how the Message Generator and the Formulator interact with gesture production.

The Communication Planner and the Discourse Model

The Communication Planner relies crucially on the Discourse Model in order to determine what information to encode, in what order, and in what modality. The Discourse Model has two subcomponents (Kita, 2010): the Interaction Record and the Addressee Model. The Interaction Record keeps track of what information has been communicated by the speaker and the communication partners. The Addressee Model specifies various properties of communication partners.

Interaction Record. Gesture production is sensitive to what has been communicated or not communicated in conversation. Based on qualitative analysis of when gestures appear in narrative, McNeill (1992) proposed that gestures tend to appear when the speech content makes a significant departure from what is taken to be given in the conversation (e.g., what has already been established in preceding discourse). Sometimes gestures explicitly encode the fact that certain information is in the Interaction Record. For example, during conversation, the speaker points to a conversational partner to indicate who has brought up a certain topic in an earlier part of the conversation (J. B. Bavelas, Chovil, Lawrie, & Wade, 1992). The Interaction Record includes not only what has been said but also what has been gestured and how gestures encoded the information. In a task in which participants describes a network of dots connected by lines, speakers sometimes produce a gesture that expresses the overall shape of the network at the beginning of a description. When such a preview gesture is produced, the verbal description of the network includes directional information less often, presumably because the initial overview gesture has already provided directional information (Melinger & Levelt, 2004). The Interaction Record also includes information about how certain information has been gesturally encoded. When the speaker gesturally express semantically related contents in different parts of a conversation, these gestures tend to share form features ("catchment" in McNeill, 2005). Similarly, when two speakers gesturally express the same referent in conversation, the handshapes of the two speakers' co-referential gestures tend to converge, but only when they can see each other (Kimbara, 2008). Thus, how the other speaker gesturally encoded a particular entity is stored in the Interaction Record and recycled in production. When the same entities are referred to repeatedly in a story, each tends to be expressed in a particular location in space (Gullberg, 2006; McNeill, 1992), not unlike anaphora in sign language (e.g., Engberg-Pedersen, 2003).

Addressee Model. Gesture production is modulated by what speakers know about the addressee. Relevant properties of the addressee include interactional potential, perceptual potential, cognitive potential, epistemic status, and attentiveness. The interaction potential refers to the degree to which the addressee can react to the speaker's utterances online, and it influences the gesture frequency. When speakers have an interactive addressee (e.g. talking on the phone), they produce gestures more frequently than when they do not (e.g., speaking to a tape recorder) (J. Bavelas, et al., 2008; Cohen, 1977). The perceptual potential of the addressee also influences the gesture frequency. Speakers produce gestures more often when the addressee can see the gestures (Alibali, Heath, & Myers, 2001; Cohen, 1977)[1]. The cognitive potential of the addressee influences the gesture frequency as well as the way in which gestures are produced. When speakers use ambiguous words (homophones, drinking "glasses" vs. optical "glasses"), they are likely to produce iconic gestures that disambiguate speech (Holler & Beattie, 2003; Kidd & Holler, 2009). Similar sensitivity to the addressee's ability to identify the referent has been shown in a corpus analysis of naturalistic data (Enfield, Kita, & de Ruiter, 2007). In the corpus, speakers are describing how their village and its surrounding area have changed to somebody who

[1] The Communication Planner has to also obtain information from the Environment to assess perceptual accessibility of gestures.

is not as knowledgeable about the area. Small pointing gestures often accompanied verbal expression of landmarks when it is likely but not certain that the referent can be identified by the addressee. The addressee's epistemic state, namely what the addressee knows, also influences the way gestures are produced. When the speaker talks about things for which the speaker and addressee have shared information, gestures tend to be less precise (Gerwing & Bavelas, 2004) though shared knowledge mostly do not make gestures less informative (Holler & Wilkin, 2009). Finally, the listener's attention state modulates the frequency of gestures. Speakers produce gestures more frequently when the addressee is attending to the speakers than when s/he is not (Jacobs & Garnham, 2007).

The Communication Planner and the Environment

One of the tasks of the Communication Planner is to decide roughly what information will be conveyed in what modality. This may depend on the properties of the Environment in which communication takes place. For example, imagine a referential communication task in which two participants, the director and the matcher, are seated side by side in front of an array of photographs of faces. The director describes one of the photographs, and the matcher has to identify which photograph is the referent. In this situation, participants use pointing gestures with deictic expressions such as *this* and *here* to identify a photograph more often when the participants are close to the array (arms length or 25 cm) than when they are further away (50 cm – 100 cm) (Bangerter, 2004). Conversely, the participants use verbal expressions to identify a photograph less often when the array is close because gestures can fully identify the referent. That is, depending on the distance to the referent, speakers distribute information differently between the gesture and speech modalities in order to optimise communication (see also van der Sluis & Krahmer, 2007 for similar results).

The Action Generator and the physical or imagined Environment

The gesture production process needs access to the information about the Environment for various reasons. This is necessary when gestures need to take into account physical obstacles (e.g., so as not to hit the listener) (de Ruiter, 2000) or when producing gestures that point at or trace a physically present target. Sometimes, gestural depiction relies on physical props. For example, a pointing gesture that indicates a horizontal direction and comes to contact with a vertical piece of timber in a door frame may depict a contraption with a horizontal bar supported by two vertical poles (Haviland, 2003). In this example, the vertical piece of timber represents the vertical poles. Production of such a gesture requires representation of the physical environment.

Gestures can be produced within an imagined environment that is generated on the basis of information activated in visuospatial and motoric working memory. Gestures are often produced as if there are imaginary objects (e.g., a gesture that depicts grasping of a cup). Gestures can take an active role in establishing and enriching the imagined environment (McNeill, 2003); that is, gestures can assign meaning to a specific location in the gesture space ("abstract deixis", McNeill, 1992; McNeill, Cassell, & Levy, 1993). The boundary between the physical and imagined environments is not clear-cut. For example, gestures can be produced near a physically present object in order to depict an imaginary transformation of the object. When describing how a geometric figure on a computer screen can be rotated, participants often produce gestures near the computer screen, as if the hand grasps the object and rotates it (Chu & Kita, 2008) (see also LeBaron & Streeck, 2000).

Interaction between the Message Generator and the Action Generator

Speech-to-gesture influence: syntactic packaging. The speech production process can influence the gesture production process via the link between the Message Generator and the Action Generator. The Message Generator creates the propositional content for utterances. Given the evidence that a clause (a grammatical unit controlled by a verb) is an important planning unit for speech production (Bock & Cutting, 1992), it can be assumed that the Message Generator packages information that is readily verbalizable within a clause. The way speech packages information is reflected in the gestural packaging information, as demonstrated by studies summarized below.

The speech-gesture convergence in information packaging can be demonstrated in the domain of motion events. Languages vary as to syntactic packaging of information about manner (how something moves) and path (which direction something moves). Some languages (e.g., English) typically encode manner and path within a single clause (e.g., "he rolled down the hill"), while others (e.g., Japanese and Turkish) typically use two clauses (e.g., "he descended the hill, as he rolled"). When describing motion events with manner and path, English speakers are more likely to produce a single gesture that encoding both manner and path (e.g., a hand traces a circular movement as it moves across in front of the torso). In contrast, Japanese and Turkish speakers are more likely to produce two separate gestures for manner and path (Kita & Özyürek, 2003; Özyürek et al., 2008). The same effect can be shown within English speakers. One-clause and two-clause descriptions can be elicited from English speakers, using the following principle. When the strength of the causal link between manner and path (e.g., whether rolling causes descending) is weaker, English speakers tend to deviate from the typical one-clause description and increase the use of two-clause descriptions similar to Turkish and Japanese (Goldberg, 1997). English speakers tend to produce a single gesture encoding both manner and path when they encode manner and path in a single clause, but produce separate gestures for manner and path when they encode manner and path in two different clauses (Kita et al., 2007). Finally, the link between syntactic packaging in speech and gesture can also be seen in Turkish learners of English at different proficiency levels. Turkish speakers who speak English well enough to package manner and path in a single clause tend to produce a gesture that encodes both manner and path. In contrast, Turkish speakers whose proficiency level is such that they still produce two-clause descriptions in English (presumably transfer from Turkish) tend to produce separate gestures for manner and path (Özyürek, 2002)

Speech-gesture influence: conceptualization load. In line with the idea that gesturing facilitates conceptualisation for speaking, gesture frequency increases when the conceptualisation load is higher (Hostetter, Alibali, & Kita, 2007b; Kita & Davies, 2009; Melinger & Kita, 2007). For example (Figure 2), imagine the situation in which participants are instructed to describe the content of each of the six rectangles, while ignoring the difference between the dark versus light coloured lines. The dark lines disrupt how information should be packaged in the hard condition (e.g., in the left top rectangle in Figure 2 (b), it is difficult to conceptualize the entire diagonal line as a unit for verbalization), but not in the easy condition. Speakers produce more representational gestures in the hard condition than in the easy condition. When it is more difficult to package information into units for speech production, that is, when conceptualisation for speaking (in particular, microplanning in Levelt, 1989) is more difficult, gesture production is triggered.

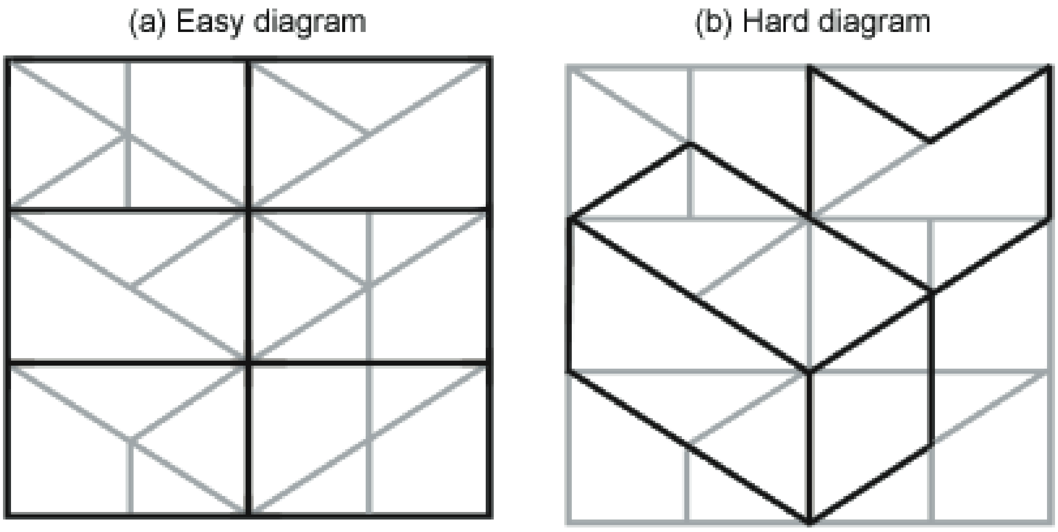

Figure 2. Example of a stimulus pair that manipulate conceptualisation load during description (Kita & Davies, 2009). (permission pending).

Gesture-to-speech influence. The gesture production process can influence the speech production process via the link from the Action Generator to the Message Generator. The nature of this link has been investigated in studies that manipulated how and whether gesture is produced, as summarised below.

How information is grouped into gestures shapes how the same information is grammatically grouped in speech. When Dutch speakers describe motion events with manner and path components (e.g., rolling up), the type of gestures they are instructed to produce influence the type of grammatical structures (Mol & Kita, in press). When the speakers are instructed to produce a single gesture encoding both manner and path, they are more likely to linguistically package manner and path in a single clause (e.g., "he rolls upwards"), but when they produced two separate gestures for manner and path, they are more likely to distribute manner and path expressions across two clauses (e.g., "he turns as he goes up"). In other words, what is encoded in a gesture is likely to be linguistically expressed within a clause, which is an important speech-planning unit (Bock & Cutting, 1992).

The information highlighted by gestures is fed into the Message Generator and is likely to be verbally expressed (Alibali & Kita, 2010; Alibali, Spencer, Knox, & Kita, 2011). When five to seven year old children are asked to explain answers to a Piagetian conservation task, the content of their explanation varied as a function of whether or not they were allowed to gesture (Alibali & Kita, 2010). In a Piagetian conservation task, the children are presented with two entities with the identical quantity (e.g., two identical glasses with water up to the same level). Then, the experimenter transforms the appearance of one entity in front of the child (e.g., pours water from one of the glasses into a wider and shallower dish) and asks which entity has more water. Five to seven year old children find this task difficult (e.g., they tend to think that there is more water in the thinner and taller glass than in the wider and shallower dish). After children have answered the quantity question, the experimenter asks the reason for their answer. When children are allowed to gesture, they tend to gesture about various aspects of the task objects (e.g., width or height of a glass). Crucially, children's explanations include features of task objects in front of them (e.g., "because this one is tall and that one is short") more often when they are allowed to gesture than when they are not. That is, when gesture highlight certain information, the information is likely to be included in the message that speakers generate for their explanation (see also Alibali & Kita, 2011). That is, gesture influences "microplanning" (Levelt, 1989) in the conceptualisation process, in which a message for each utterance is determined.

Manipulation of gestures influences fluency of speech production. When speakers describe spatial contents of an animated cartoon, the speech rate is higher and disfluencies are less frequent when the speakers are allowed to gesture than when they are prohibited from gesturing (Rauscher, Krauss, & Chen, 1996). This is compatible with the idea that gesture facilitates verbal encoding of spatial information.

The exact nature of the gestural influence on speech production is much debated in the literature. There are three views, which are not mutually exclusive. The first view is that gesture facilitates conceptualisation for speaking (Kita, 2000) which is compatible with the model in Figure 1. There is substantial evidence for this view (Alibali & Kita, 2010; Alibali, Kita, Bigelow, Wolfman, & Klein, 2001; Alibali, Kita, & Young, 2000; Alibali, et al., 2011; Hostetter, Alibali, & Kita, 2007a; Hostetter, et al., 2007b; Kita, 2000; Kita & Davies, 2009; Melinger & Kita, 2007; Mol & Kita, in press). The second view is that gesture facilitates lexical retrieval (Krauss, Chen, & Gottesman, 2000; Rauscher, et al., 1996). There is very limited evidence that uniquely supports this hypothesis (see Beattie & Coughlan, 1999 for further discussions; Kita, 2000) (but see Rose, 2006). The third view is that gesture activates imagery whose content is to be verbally expressed (Bock & Cutting, 1992; de Ruiter, 1998; Wesp, Hesse, Keutmann, & Wheaton, 2001). The evidence for this view is that speakers produce more gestures when they have to describe stimuli from memory than when they can see the stimuli during description. In the memory condition, the image of the visual stimuli needs to be activated and, presumably, more gestures are produced in order to activate the necessary images. However, there is no study supporting this view that manipulated availability of gestures.

Other models of speech-gesture production

This article used Kita and Özyürek's (2003) model to summarise what is known about production of speech-accompanying gestures. However, it is important to acknowledge that there are other models. De Ruiter's (2000) model and Krauss and his colleagues' (2000) model are also based on Levelt's (1989) model of speech production. These models differ from the model in Figure 1 in the way gestural contents are determined. The content of gesture is determined by the conceptual planning process (the Conceptualizer in Levelt, 1989) in de Ruiter (2000) but in spatial working memory in Krauss et al. (2000). Unlike Figure 1, both models do not allow feedback from the formulation process to the conceptualisation process. Consequently, they cannot account for the findings that syntactic packaging of information influencing gestures.

It is also important to note theories of speech-gesture production that do not use the box-and-arrow architecture. Growth Point theory (McNeill, 1985, 1992, 2005; McNeill & Duncan, 2000) is very influential in its claim that speech and gesture production are an integrated process (see also Kendon, 1980). This theory brought gesture into psycholinguistics. According to the Growth Point theory, the information that stands out from the context forms a "Growth Point", which has both imagistic and verbal aspects. The imagistic aspect develops into a gesture and the verbal aspect develops into speech that is semantically associated with the gesture. Another more recent theory is the Gesture as a Simulated Action theory (Hostetter & Alibali, 2010). This theory assumes that underlying semantic representation for speech is motor or perceptual simulation (Barsalou, 1999) and gestures are generated from the same motor or perceptual simulation. When the strength of a simulation exceeds a certain threshold, a gesture is produced.

Other important issues

Due to space limitations, this article did not cover the following issues relevant to the relationship between speech and gesture production. The first issue is cultural variation in gesture production and reasons for the variation (Kita, 2009). The second issue is the model for how speech and gesture are synchronised. Most of the work on synchronisation is on pointing gestures (de Ruiter, 1998; Levelt, Richardson, & La Heij, 1985). Representational gestures tend to precede co-expressive words (McNeill, 1992; Morrel-Samuels & Krauss, 1992); however, the mechanism for this synchronisation has not been clarified. The third issue is how the relationship between speech and gesture production develops during childhood (Capirci, et al., 1996; Iverson & Goldin-Meadow, 2005; Nicoladis, 2002; Nicoladis, Mayberry, & Genesee, 1999; Özyürek, et al., 2008; S. Stefanini, Bello, Caselli, Iverson, & Volterra, 2009). The fourth issue is the neural substrates for the production of speech-accompanying gestures (Cocks, Dipper, Middleton, & Morgan, 2011; Hadar, Burstein, Krauss, & Soroker, 1998; Hadar & Krauss, 1999; Hadar, Wenkert-Olenik, Krauss, & Soroker, 1998; Hadar & Yadlin-Gedassy, 1994; Hogrefe, Ziegler, Tillmann, & Goldenberg, in press; Kimura, 1973a, 1973b; Kita, de Condappa, & Mohr, 2007; Kita & Lausberg, 2008; Lausberg, Davis, & Rothenhäuser, 2000; Rose, 2006). The fifth issue is how gesture production is affected in developmental disorders such as Specific Language Impairment, autism, Down syndrome and Williams syndrome (Bello, Capirci, & Volterra, 2004; de marchena & Eigsti, 2010; Evans, Alibali, & McNeil, 2001; Silvia Stefanini, Caselli, & Volterra, 2007; Volterra, Capirci, & Caselli, 2001).

Conclusion

Speech-accompanying gestures are tightly coordinated with speech production. Gesture and speech are planned together as an integrated communicative move (Kendon, 2004). What is expressed in gesture and how it is expressed are shaped by information in the physical environment, discursive contexts, and how speech formulates information to be conveyed. Thus, it is not sufficient just to observe speech production to fully understand human communication.

References

Alibali, M. W., Heath, D. C., & Myers, H. J. (2001). Effects of visibility between speaker and listener on gesture production: some gestures are meant to be seen. *Journal of Memory and Language, 44,* 169-188.

Alibali, M. W., & Kita, S. (2010). Gesture highlights perceptually present information for speakers. *Gesture, 10*(1), 3-28.

Alibali, M. W., Kita, S., Bigelow, L. J., Wolfman, C. M., & Klein, S. M. (2001). Gesture plays a role in thinking for speaking. In C. Cavé, I. Guaïtella & S. Santi (Eds.), *Oralité et gesturalité: Interactions et comportements multimodaux dans la communication* (pp. 407-410). Paris: L'Harmattan.

Alibali, M. W., Kita, S., & Young, A. J. (2000). Gesture and the process of speech production: we think, therefore we gesture. *Language and Cognitive Processes, 15,* 593-613.

Alibali, M. W., Spencer, R. C., Knox, L., & Kita, S. (2011). Spontaneous Gestures Influence Strategy Choices in Problem Solving. *Psychological Science, 22*(9), 1138-1144.

Bangerter, A. (2004). Using pointing and describing to achieve joint focus of attention in dialogue. [Article]. *Psychological Science, 15*(6), 415-419.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral And Brain Sciences, 22*(4), 577-680.

Bavelas, J., Gerwing, J., Sutton, C., & Prevost, D. (2008). Gesturing on the telephone: Independent effects of dialogue and visibility. *Journal of Memory and Language, 58*(2), 495-520.

Bavelas, J. B., Chovil, N., Lawrie, D., A., & Wade, A. (1992). Interactive gestures. *Discourse Processes, 15,* 269-189.

Beattie, G., & Coughlan, J. (1999). An experimental investigation of the role of iconic gestures in lexical access using the tip-of-the-tongue phenomenon. *British journal of psychology, 90,* 35-56.

Bello, A., Capirci, O., & Volterra, V. (2004). Lexical production in children with Williams syndrome: spontaneous use of gesture in a naming task. *Neuropsychologia, 42*(2), 201-213.

Bock, K., & Cutting, J. C. (1992). Regulating Mental Energy: Performance Units in Language Production. *Journal of Memory and Language, 31*, 99-127.

Capirci, O., Iverson, J. M., Pizzuto, E., & Volterra, V. (1996). Gestures and words during the transition to two-word speech. *Journal of Child Language, 23*(3), 645-673.
Chu, M., & Kita, S. (2008). Spontaneous gestures during mental rotation tasks: Insights into the microdevelopment of the motor strategy. *Journal of Experimental Psychology: General, 137*, 706-723.
Cienki, A., & Müller, C. (Eds.). (2008). *Metaphor and gesture*. Amsterdam: John Benjamins.
Cocks, N., Dipper, L., Middleton, R., & Morgan, G. (2011). The impact of aphasia on gesture production: A case of condution aphasia. *International Journal of Language and Communication Disorders, 46*(4), 423-436.
Cohen, A. A. (1977). The communicative functions of hand illustrators. *Journal of Communication, 27*(4), 54-63.
de marchena, A., & Eigsti, I. M. (2010). Conversational gestures in autism spectrum disorders: Asycnrhony but not decreased frequency. *Autism Research, 3*, 311-322.
de Ruiter, J. P. (1998). *Gesture and speech production.* Doctoral dissertation. University of Nijmegen, Nijmegen, the Netherlands.
de Ruiter, J. P. (2000). The production of gesture and speech. In D. McNeill (Ed.), *Language and gesture* (pp. 284-311). Chicago: University of Chicago Press.
Enfield, N. J., Kita, S., & de Ruiter, J. P. (2007). Primary and secondary pragmatic functions of pointing gestures. *Journal of Pragmatics, 39*, 1722-1741.
Engberg-Pedersen, E. (2003). From pointing to reference and predication: Pointing signs, eyegaze, and head and body orientation in Danish Sign Language. In S. Kita (Ed.), *Pointing: Where language, cognition and culture meet* (pp. 269-292). Mahwah, NJ: Lawrence Erlbaum.
Evans, J. L., Alibali, M. W., & McNeil, N. M. (2001). Divergence of verbal expression and embodied knowledge: Evidence from speech and gesture in children with specific language impairment? *Language and Cognitive Processes, 16*(2-3), 309-331.
Gerwing, J., & Bavelas, J. (2004). Linguistic influences on gesture's form. *Gesture, 4*(2), 157-195.
Goldberg, A. E. (1997). The relationship between verbs and constructions. In M. Verspoor, K. D. Lee & E. Seetser (Eds.), *Lexical and syntactical constructions and the constructionof meaning* (pp. 383-398). Amsterdam: John Benjamins.
Gullberg, M. (2006). Handling discourse: Gestures, reference tracking, and communication strategies in early L2. *Language Learning, 56*(1), 155-196.
Hadar, U., Burstein, A., Krauss, R., & Soroker, N. (1998). Ideational gestures and speech in brain-damaged subjects. *Language and Cognitive Processes, 13*(1), 59-76.
Hadar, U., & Krauss, R. K. (1999). Iconic gestures: the grammatical categories of lexical affilates. *Journal of Neurolinguistics, 12*, 1-12.
Hadar, U., Wenkert-Olenik, Krauss, R., & Soroker, N. (1998). Gesture and the processing of speech: neuropsychological evidence. *Brain and Language, 62*, 107-126.
Hadar, U., & Yadlin-Gedassy, S. (1994). Conceptual and lexical aspects of gesture: evidence from aphasia. *Journal of Neurolinguistics, 8*, 57-65.
Haviland, J. B. (2003). How to point in Zinacantán. In S. Kita (Ed.), *Pointing: Where language, cognition, and culture meet* (pp. 139-169). Mahwah, NJ: Lawrence Erlbaum.

Hogrefe, K., Ziegler, W., Tillmann, C., & Goldenberg, G. (in press). Non-verbal communication in severe aphasia: Influence of aphasia, apraxia, or semantic processing? *Cortex*.

Holler, J., & Beattie, G. (2003). Pragmatic aspects of re- presentational gestures: do speakers use them to clarify ver- bal ambiguity with the listener? . *Gesture, 3*, 127-154.

Holler, J., & Wilkin, K. (2009). Communicating common ground: How mutually shared knowledge influences speech and gesture in a narrative task. *Language and Cognitive Processes, 24*(2), 267-289.

Hostetter, A. B., & Alibali, M. W. (2010). Language, gesture, action! A test of the Gesture as Simulated Action framework. *Journal of Memory and Language, 63*(2), 245-257.

Hostetter, A. B., Alibali, M. W., & Kita, S. (2007a). Does sitting on your hands make you bite your tongue? The effects of gesture prohibition on speech during motor description. In D. S. McNamara & J. G. Trafton (Eds.), *Proceedings of the twenty ninth annual conference of the Cognitive Science Society* (pp. 1097-1102). Mahwah, NJ: Lawrence Erlbaum.

Hostetter, A. B., Alibali, M. W., & Kita, S. (2007b). I see it in my hand's eye: Representational gestures are sensitive to conceptual demands. *Language and Cognitive Processes, 22*(3), 313-336.

Iverson, J. M., & Goldin-Meadow, S. (2001). The resilience of gesture in talk: gesture in blind speakers and listeners. *Developmental Science, 4*(4), 416-422.

Iverson, J. M., & Goldin-Meadow, S. (2005). Gesture paves the way for language development. *Psychological Science, 16*(5), 367-371.

Jacobs, N., & Garnham, A. (2007). The role of conversational hand gestures in a narrative task. *Journal of Memory and Language, 56*(2), 291-303.

Kendon, A. (1980). Gesticulation and speech: Two aspects of the process of utterance. In M. R. Key (Ed.), *The relation between verbal and nonverbal communication* (pp. 207-227). The Hague: Mouton.

Kendon, A. (1992). Some Recent Work from Italy on Quotable Gestures (Emblems). *Journal of Linguistic Anthropology, 2*(1), 92-108.

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

Kidd, E., & Holler, J. (2009). Children's use of gesture to resolve lexical ambiguity. *Developmental Science, 12*(6), 903-913.

Kimbara, I. (2008). Gesture form convergence in joint description. *Journal of Nonverbal Behavior, 32*(2), 123-131.

Kimura, D. (1973a). Manual activity during speaking - i. Right-handers. *Neuropsychologia, 11*, 45-50.

Kimura, D. (1973b). Manual activity during speaking- II. Left-handers. *Neuropsychologia, 11*, 51-55.

Kita, S. (2000). How representational gestures help speaking. In D. McNeill (Ed.), *Language and gesture* (pp. 162-185). Cambridge: Cambridge University Press.

Kita, S. (2003). *Pointing: where language, culture, and cognition meet*. Mahwah, NJ: Lawrence Erlbaum.

Kita, S. (2009). Cross-cultural variation of speech-accompanying gesture: A review. *Language and Cognitive Processes, 24*(2), 145-167.

Kita, S. (2010). A model of speech production. In E. Morsella (Ed.), *Expressing onself / expressing one's self: Communication, cognition, language, and identity* (pp. 9-22). New York, London: Psychology Press.

Kita, S., & Davies, T. S. (2009). Competing conceptual representations trigger co-speech representational gestures. *Language and Cognitive Processes, 24*(5), 761-775.
Kita, S., de Condappa, O., & Mohr, C. (2007). Metaphor explanation attenuates the right-hand preference for depictive co-speech gestures that imitate actions. *Brain and Language*, *101*(3), 185.
Kita, S., & Lausberg, H. (2008). Generation of co-speech gestures based on spatial imagery from the right hemisphere: Evidence from split-brain patients. *Cortex, 44*, 131-139.
Kita, S., & Özyürek, A. (2003). What does cross-linguistic variation in semantic coordination of speech and gesture reveal?: Evidence for an interface representation of spatial thinking and speaking. *Journal of Memory and Language, 48*, 16-32.
Kita, S., Özyürek, A., Allen, S., Brown, A., Furman, R., & Ishizuka, T. (2007). How do our hands speak? Mechanisms underlying linguistic effects on representational gestures. *Language and Cognitive Processes, 22*(8), 1-25.
Krahmer, E., & Swerts, M. (2007). Effect of visual beats on prosodic prominence: Acoustic analyses, auditory perception, and visual perception. *Journal of Memory and Language, 57*, 396-414.
Krauss, R. M., Chen, Y., & Gottesman, R. F. (2000). Lexical gestures and lexical access: a process model. In D. McNeill (Ed.), *Language and gesture* (pp. 261-283). Cambridge: Cambridge University Press.
Lausberg, H., Davis, M., & Rothenhäuser, A. (2000). Hemispheric specialization in spontaneous gesticulation in a patient with callosal disconnection. *Neuropsychologia, 38*, 1654-1663.
LeBaron, C. D., & Streeck, J. (2000). Gesture, knowledge, and the world. In D. McNeill (Ed.), *Gesture and language* (pp. 118-138). Chicago: University of Chicago Press.
Levelt, W. J. M. (1989). *Speaking*. Cambridge, MA: The MIT Press.
Levelt, W. J. M., Richardson, G., & La Heij, W. (1985). Pointing and voicing in deictic expressions. *Journal of memory and language, 24*, 133-164.
McNeill, D. (1985). So you think gestures are nonverbal. *Psychological Review, 92*, 350-371.
McNeill, D. (1992). *Hand and mind*. Chicago: University of Chicago Press.
McNeill, D. (2003). Pointing and morality in Chicago. In S. Kita (Ed.), *Pointing: Where language, cognition, and culture meet.* (pp. 293-306). Mahwah, NJ: Lawrence Erlbaum.
McNeill, D. (2005). *Gesture and thought*. Chicago: University of Chicago Press.
McNeill, D., Cassell, J., & Levy, E. T. (1993). Abstract deixis. *Semiotica, 95*(1-2), 5-19.
McNeill, D., & Duncan, S. D. (2000). Growth points in thinking-for-speaking. In D. McNeill (Ed.), *Language and gesture* (pp. 141-161). Cambridge: Cambridge University Press.
Melinger, A., & Kita, S. (2007). Conceptualization load triggers gesture production. *Language and Cognitive Processes, 22*(4), 473-500.
Melinger, A., & Levelt, W. J. M. (2004). Gesture and the communicative intention of the speaker. *Gesture, 4*(2), 119-141.
Mol, L., & Kita, S. (in press). Gesture structure affects syntactic structure in speech *Proceedings of the thirty first annual conference of the Cognitive Science Society* Austin, TX: Cognitive Science Society.

Morrel-Samuels, P., & Krauss, R. M. (1992). Word familiarity predicts temporal asynchrony of hand gestures and speech. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 615-622.

Morris, D., Collett, P., Marsh, P., & O'Shaughnessy, M. (1979). *Gestures, their origins and distribution*. New York: Stein and Day.

Nicoladis, E. (2002). Some gestures develop in conjunction with spoken language development and others don't: Evidence from bilingual preschoolers. *Journal of Nonverbal Behavior, 26*(4), 241-266.

Nicoladis, E., Mayberry, R. I., & Genesee, F. (1999). Gesture and early bilingual development. *Developmental psychology, 35*(2), 514-526.

Özyürek, A. (2002). Speech-gesture synchrony in typologically differnt languages and second language acquistion. In B. Skarabela, S. Fish & A. H. J. Do (Eds.), *Proceedings from the 26th Annual Boston University Conference in Language Development* (pp. 500-509). Somerville, MA: Cascadilla Press.

Özyürek, A., Kita, S., Allen, S., Brown, A., Furman, R., & Ishizuka, T. (2008). Development of cross-linguistic variation in speech and gesture: Motion events in English and Turkish. *Developmental Psychology, 44*(4), 1040-1054.

Rauscher, F. H., Krauss, R. M., & Chen, Y. (1996). Gesture, speech, and lexical access: the role of lexical movements in speech production. *Psychological Science, 7*, 226-230.

Rose, M. (2006). The utility of arm and hand gestures in the treatment of aphasia. *Advances in Speech-Language Pathology, 8*(2), 92.

Stefanini, S., Bello, A., Caselli, M. C., Iverson, J. M., & Volterra, V. (2009). Co-speech gestures in a naming task: Developmental data. *Language and Cognitive Processes, 24*(2), 168-189.

Stefanini, S., Caselli, M. C., & Volterra, V. (2007). Spoken and gestural production in a naming task by young children with Down syndrome. *Brain and Language, 101*(3), 208.

van der Sluis, I., & Krahmer, E. (2007). Generating multimodal references. *Discourse Processes, 44*, 145-174.

Volterra, V., Capirci, O., & Caselli, M. C. (2001). What atypical populations can reveal about language development: The contrast between deafness and Williams syndrome. *Language and Cognitive Processes, 16*(2-3), 219-239.

Wesp, R., Hesse, J., Keutmann, D., & Wheaton, K. (2001). Gestures maintain spatial imagery. *American Journal of Psychology, 114*, 591-600.