Does Gesture Help Processes of Speech Production? Evidence for Conceptual Level Facilitation

ALISSA MELINGER[1] and SOTARO KITA[2]
Saarland University[1] and *Max Planck Institute for Psycholinguistics*[2]

0. Introduction

When people speak, they gesture, and these gestures are often transparently related to the semantic content expressed in the speech. Gestures that are semantically co-expressive with the co-temporal speech are called representational gestures. Representational gestures can encode various semantic features of objects and events, such as the shape of an object, the interactive characteristic of an object, the function of an object, an activity, an entity's spatial location, etc. While many have claimed that representational gestures serve a communicative function (see Kendon 1994 for a survey), it has also been observed that gestures are produced even in the absence of a visible interlocutor. When people speak on the phone or to a non-visible companion, they continue to produce co-speech gestures (Cohen 1977). This suggests that these gestures may serve a(n additional) purpose that is not communicative. Some researchers have proposed that gestures might actually aid processes of speech production; however, there is a great deal of disagreement about how this might work (see Kita 2000 for a review of the debate). This paper investigates the issue of how gestures aid speech production processes.

1. Stages of Speech Production

It is generally agreed that processes of speech production are incremental and can be broken down into three processing stages: (a) conceptualization, (b) formulation, and (c) articulation (Levelt 1989). According to Levelt, at the conceptualization stage the speaker conceives of an intention, selects the relevant information to be expressed, and orders the information for expression, among other processes. It is within the conceptual level that "thinking for speaking", as intended by Slobin (1987), occurs. At the formulation stage the pre-linguistic message is given linguistic form. Lexical items and syntactic frames are selected, ordered, and combined at this stage. Finally, at the articulation stage signals are transmitted to the articulators to produce the desired utterance.

Additionally, there is a large body of evidence from several distinct sources suggesting that the formulation stage is further divided into two distinct levels of

representation in which different types of information are retrieved. First, an abstract lexico-semantic/syntactic representation (or lemma, cf. Kempen and Huijbers 1983) is retrieved, based on the semantic information passed down from the conceptualizer. In a separate stage of processing, the form of the word is specified. At this word form level, the metrical, segmental, and morphological structure of the word is specified.

2. What Role Does Gesture Play for the Speaker?

Many researchers have speculated on how gestures might contribute to the general processes involved in speech production. Many researchers propose that gesture facilitates speaking by aiding in the process of lexical retrieval, i.e. during formulation. We will refer to this hypothesis as the Lexical Level Hypothesis. Specifically, the claim is that a gesture serves as a cross-modal prime to boost the activation of a particular lexical entry, either at the lemma level (Krauss et al. 1996) or the word form level (Butterworth and Hadar 1989, Krauss et al. 2000). They cite the fact that when gesture production is restricted, speech including spatial content is adversely affected (Rauscher et al. 1996). Furthermore, gesture prohibition has also been shown to increase the number of retrieval failures in a tip-of-the-tongue elicitation study (Frick-Horbury and Guttentag 1998). Proponents of the Lexical Level Hypothesis interpret these results as evidence that gesturing facilitates lexical retrieval.

Another interpretation of these same data, however, is that gesture prohibition adversely affects the conceptual processes involved in constructing a pre-linguistic message rather than in lexical retrieval processes per se. If gestures aid in the activation of conceptual representations, this activation will then spread to the lexical level. Thus, effects such as those used to support the Lexical Level Hypothesis are also consistent with a model in which gestures aid processes at the conceptual level. Supporters of this view argue that gestures help activate imagistic and conceptual information at a pre-lexical level (de Ruiter 1998) and help to map between imagistic information and propositional information (Alibali et al. 2000). We will refer to the hypothesis that gesture facilitates conceptual level processes as the Conceptual Level Hypothesis.

One problem with both the lexical and conceptual level hypotheses is that certain details crucial to formulating testable predictions are underspecified. For example, very little is known about the general conceptual processes that are involved with speaking, and therefore it is difficult to be explicit about how gesture could facilitate these processes. On the other hand, it is also difficult to understand exactly how a gesture could prime lexical information at, for example, the word form level, given that evidence from speech production literature suggests that semantic information is not represented at that level. In other words, why should making a circular motion with a finger help the retrieval of a representation that specifies the number of syllables or the segmental content for the word circumference? It seems more likely that a gesture should prime semantic representations rather than phonological representations. However, even claiming that gestures

prime semantic representations is compatible with both hypotheses because there is a great deal of debate in the production literature as to whether lemmas contain any semantic information themselves (Butterworth 1989) or whether they only have meaning by virtue of connections to the conceptual level (Levelt et al. 1999, Dell 1986). Furthermore, given that lemmas are activated as the result of activation passed on from the conceptual level, it is quite difficult to distinguish effects that occur as a direct link between gesture and the lexicon from effects that have their locus at the conceptual level and then trickle down to the lexical level. In fact, these two possibilities are so indistinguishable that most researchers who propose the latter must also allow for the possibility of the former (Krauss et al. 1996, Krauss et al. 2000).

Despite the difficulties in distinguishing these two hypotheses, they do predict a different distribution of gesture-speech interactions; they predict that linguistic differences will have an effect on gesture under very different circumstances. For example, the Lexical Level Hypothesis predicts that gestures are produced when lexical retrieval is more difficult due to inherent characteristics of the target word, such as having low frequency or many lexical competitors, or when retrieval is more difficult due to external or contextual factors. For example, if a word is produced twice in immediately adjacent clauses, the second mention should be easier to retrieve then the first mention, since the lemma should have some residual activation from when the word was first produced. The relative ease of producing the second mention should result in fewer co-expressive gestures then were produced for the first mention if gestures are produced to facilitate lexical retrieval. In contrast, the Conceptual Level Hypothesis predicts no effect of lexical pre-activation. The Conceptual Level Hypothesis predicts that gestures are produced when the mapping from imagistic representations to propositional representations is complicated or when the information at the conceptual level requires additional computations. The Lexical Level Hypothesis predicts no effect of conceptual complexity.

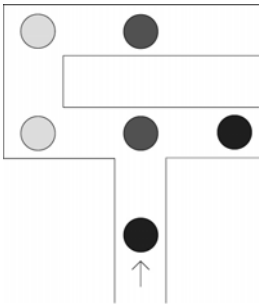
To test these hypotheses, we analyzed a series of brief picture descriptions and compared when gestures were produced to when gestures were not produced. We then examined the results with respect to the predictions of these two hypotheses.

3. Experiment

We presented 16 native speakers of Dutch with 16 abstract map-like images. The maps depicted segments of streets and intersections with destinations arranged along a path. Destinations were large colored circles positioned in the middle of the streets.[1] Images contained either five or six destinations, and half included a branching route. An example of the images used in this study is provided in (1).

[1] To facilitate memorization, destinations were limited to 3 colors: yellow, blue, and red.

(1) Sample image used to elicit path descriptions



Participants were seated across from their interlocutor with a visual block separating them. Participants were instructed to describe, from memory, a route that led past all the destinations in each image. Their description was supposed to be detailed enough to allow the non-visible listener to envision the picture and answer questions about it. Each picture was placed on a table in front of the participant for approximately 15 seconds. After this memorization period, the picture was removed and the participant began to describe the image. Participants were free to describe the routes in any way that was natural to them. They were not given any linguistic examples to bias their description strategy. The listener (who was in fact a conspirator) was instructed not to ask any specific questions that might bias the content of the descriptions. She was free, however, to ask the speaker to repeat portions or even the entire description of an image. Participants were video recorded from two angles to capture their gestural and linguistic behavior. The first and last four pictures for each participant were coded for speech and gestures.

3.1. Linguistic Data

For the purposes of conducting a quantitative analysis, several types of linguistic information were identified from the transcriptions of the participants' descriptions: directional information, destination information, landmark information, and overview information. To clarify these distinctions, consider the constructed sample description provided in (2) below. Content specifying a direction of movement, such the italicized portions of (2a) and (2d), was classified as directional information. References to the colored circles as the destination of a movement, such as the underlined portions of (2a) and (2d), were classified as destination information. References to circles as the goal of a movement were distinguished from references to circles as the source of a movement, such as the italicized portion of (2f). References to source information, as well as references to the characteristics of the road, as in the italicized portions of (2b) and (2e), were classified as landmarks, since they functioned as reference points within the image. Information about the general shape of the image or a sub-portion of the image,

such as in (2d), and information about the relationship between multiple destinations, such as in (2g), were both classified as overview information.

(2) a. *You walk straight ahead* and you come across a yellow circle.
b. The yellow circle is in the center of *an intersection*,
c. just a plus-form-like intersection.
d. When *you turn left* you come across a red circle.
e. Now you go back to the center of *the intersection*, so the yellow circle.
f. *From the yellow circle* we walk straight ahead and you see a blue circle.
g. So you have three circles all in a row.

To successfully complete the task, participants needed to minimally include direction and destination information. However, since they were given no examples of how to conduct their descriptions, they were free to include as much or as little information as they felt was needed. Thus, there was a wide amount of variation in the frequency with which non-essential landmark and overview information was included. Five participants generally included only the essential information. Five participants provided some landmarks in addition to the essential information. Six participants regularly included all four types of information in their descriptions. The average number of linguistic mentions for each of these types of linguistic information is provided in (3).

(3) Average number (per picture) of linguistic references to directions, balls, and non-essential information

	Essential information		Non-essential information	
	Directions	Destinations	Landmarks	Overview
Participants	5.1	6.6	1.5	0.3

Speakers were also free to adopt whatever spatial perspective they chose. They could describe the directions as though they were moving through the path (intrinsic perspective) or they could describe the objective directions as they were seen on the paper (deictic perspective or 'bird's eye view', cf. Levelt 1987). Deictic speakers use directional terms like up, down, left, and right. In contrast, intrinsic speakers use straight ahead, left, and right. Since intrinsic speakers move though the image, what they see on the paper and what they say often conflict; if they have traveled in a circle, left may be right and right may be left. This mismatch between what is on the picture and what is said does not occur for deictic speakers. The constructed examples in (4a) and (4b), which are descriptions of a portion of the image in (1), demonstrate the differences between these two perspectives. We will return to this issue of perspective-taking later in the paper.

(4) a. Deictic Perspective: You start at a blue circle, then you go up to a red circle. Next you go to the right and you see another blue circle. Now go back to the red circle where you just were and go to the left. You'll see a yellow circle. At the yellow circle, go up…

b. Intrinsic Perspective: You start at a blue circle, then you go straight ahead until you come to a red circle. Next you go to the right and you see another blue circle. Now, go back to the red circle where you just were and then go straight ahead until you come to a yellow circle. At the yellow circle, you go to the right…

3.2. Gestural Behavior

Using the video recordings of the picture descriptions, we identified all gestures that were semantically co-expressive with the concurrent speech. For example, if the speaker said "you go to the right" and simultaneously pointed to the right, we classified that as a co-expressive gesture. Gestures were generally produced with one or both hands. They were generally small movements produced either in the participant's lap or in front of the torso, close to the body. In addition to gestures produced with one or both hands, we also identified head movements that were co-expressive with the ongoing speech.

We found several different types of gestures, most corresponding to the different categories of linguistic information. The most common type of gestures were hand or head movements indicating the direction of movement in the description. Right and left were gestured most often, but up, down, and straight ahead also received many co-expressive gestures. The next most frequent type of gesture was the pointing gesture, which indicates the location of a destination point in the imaginary image created by the speaker in the gesture space. Sometimes, in addition to a pointing gesture in which the participant would draw a circle in the gesture space to indicate the colored circles that represented the destinations in the pictures. Other representational gestures (often produced as air drawings) depicted the overall shape of the image (e.g. an F-form), some component of the image (e.g. a T-intersection), or the relationship between two circles (e.g. one is directly above the other). The table in (5) shows the percentage of linguistic mentions that were produced with a co-expressive gesture for each type of information.

(5) Average percent of linguistic mentions that were produced with or without a gesture for direction, destination, and non-essential information

	Direction information	Destination information	Landmarks and overview
With a gesture	25 %	8 %	3 %
Without a gesture	75 %	92 %	97 %

On average, participants made 2.2 gestures per picture. However, there was also a large amount of variation between participants. Many participants produced virtually no gestures (N=5); others produced very few gestures, namely less then one gesture per picture (N=5); and others made many gestures, namely an average of five gestures per picture (N=6).

4. The Relationship Between Gesture and Speech Production

In order to assess the stage at which speech and gesture interact, we conducted a series of analyses on the data obtained from this elicitation procedure aimed at uncovering any differences in gesture patterns. Each analysis targeted one of the predictions of the two major hypotheses under consideration: namely, the Lexical Level Hypothesis and the Conceptual Level Hypothesis. To test each hypothesis, we identified points within and between pictures as more or less challenging at either a lexical or conceptual level and then compared the number of gestures that were produced in the two cases.

Our interpretation of differences in gesture frequency is based on the assumption that when a particular process is taxed, or when processing at a given level is taxed, speakers try to reduce the processing load by producing any behavior that will ease the load on that process. Thus, if gestures function to aid lexical retrieval, differences in complexity at the conceptual level should not affect the gesture behavior. Rather, any differences in gestural behavior should be attributable to differences in lexical content or factors related to lexical retrieval. Taken the other way, differences in gesture behavior should be attributable to differences in lexical retrieval. For example, the second mention of a lexical item should receive fewer co-expressive gestures than the first mention of the same lexical item when they occur in adjacent utterances. In contrast, if gestures function to aid conceptual processing, then differences in conceptual planning or processing should affect gesture frequency, while differences in lexical content or lexical processing demands should have no effect. For example, portions of the pictures that were conceptually or computationally more complex should elicit more gestures than less challenging portions of the pictures.

4.1. Did the Chance of Producing a Gesture Increase as the Computations at the Conceptual Level Increased?

If gestures facilitate conceptual processes, then differences in processing or planning demands at the conceptual stage should correspond to differences in gesture frequency. To test this hypothesis, we identified conceptually challenging sub-portions of our pictures. Specifically, we chose the points in the pictures when the route branched in two directions. At these points, participants were required to take a number of additional steps at the conceptual level. First, the participant had to decide which path to travel first. Next, they had to remember the color of the 'choice point ball' so they could return to it. Finally, after describing the first branch, they had to return to the choice point and decide which way to go next. For intrinsic speakers, this requires recalculating the next direction since they are no

longer facing the same direction they would have been facing if they had chosen to travel the second branch first. Thus, describing the movements around a choice point is relatively taxing compared to describing the other deterministic movements within the picture.

We targeted the three movements around the choice point as conceptually most challenging: (i) the initial movement away from the choice point, (ii) the return to the choice point, and (iii) the final movement away from the choice point, after which the path is deterministic again. For each participant who produced at least one gesture in some picture, we calculated the percentage of directional gestures produced in this three-movement window compared to the percentage of directional gestures produced overall. These figures, averaged across participants, are presented in (6) below.

(6) Average gesture frequency at conceptually challenging sub-portions of descriptions compared to overall frequency

	Directional terms with a gesture	Directional terms without a gesture
At decision point	48 %	52 %
Overall	33 %	67 %

As can be seen, the percentage of directional gestures produced within the three-movement window around the choice point is considerably higher than the percentage of gestures produced overall. This suggests that gesture production increases as the conceptual difficulty of the task increases. Given that the lexical content for these moves is essentially the same as other, non-choice point moves, the differences in gesture frequency cannot be attributed to processes of lexical retrieval.

4.2. Did Intrinsic Speakers Gesture More than Deictic Speakers?

Another test of the conceptual level hypothesis is to compare the percentage of gestures produced by intrinsic speakers to the percentage of gestures produced by deictic speakers. When deictic speakers provide the direction for a movement, they must examine the image in memory and identify the direction. In contrast, since intrinsic speakers move throughout the image, the direction of each movement is dependent on where they are in the image at that moment and what direction they are facing. Thus, they must compute the direction for each turn as they come to it. We argue that this is more complex a task, and this is supported by the fact that some participants demonstrate overt problems with the task. Often participants will pause when confronted with a particularly difficult calculation. Sometimes they will even turn themselves around in the chair to help visualize which direction they would be facing if they had actually moved through the picture. Thus, at the conceptual planning level, intrinsic perspective is more difficult then deictic

perspective. However, since the task remains the same, the words used to describe the movements are essentially the same.[2] The average numbers of directional gestures produced by intrinsic and deictic speakers, respectively, are shown in (7) below. As can be seen, intrinsic speakers produce many more directional gestures, on average, than deictic speakers, again supporting the Conceptual Level Hypothesis.

(7) Average number of directional terms that were produced with and without a co-speech gesture by deictic and intrinsic speakers

	Directional referents with a gesture	Directional referents without a gesture
Deictic speakers (N=5)	9.6	32.8
Intrinsic speakers (N=11)	12.2	25.8

Although it was not the case that adopting the intrinsic perspective automatically resulted in a large number of gestures, we did find that, on average, intrinsic speakers produced more directional gestures than deictic speakers. In contrast, their average number of linguistic mentions did not differ. Thus, while the linguistic content was essentially the same, their gesture behavior was different. As with the choice point data, these results cannot easily be interpreted as lexical retrieval facilitation.

We have shown two sources of evidence that complexity at the conceptual level increases the number of gestures produced. This seems to be strong evidence to support the Conceptual Level Hypothesis. However, the Conceptual Level Hypothesis and the Lexical Level Hypothesis need not be mutually exclusive. It is possible that gestures could facilitate processes at both levels. Therefore, our next analysis aimed to find some evidence that gestures facilitate lexical retrieval.

4.3. Did the Chance of Producing a Gesture Decrease as Ease of Lexical Retrieval Increased?

If gestures aid processes of lexical retrieval, then moments of relatively difficult retrieval should produce more gestures than moments of relatively easy retrieval. To investigate this prediction, we analyzed the gesture frequency on all lexical items that were repeated in close proximity with the same intended referent.[3] Models of word production and recognition have found that the activation of a word is facilitated if it is preceded by a related word. This facilitation is greatest when the preceding word is identical to the target word. Based on this well-established finding, we assume that the second mention of any word should be easier to retrieve then the first mention of the same word when they occur within

[2] Right and left are the same. Deictic speakers use up and down, while intrinsic speakers use straight instead.

[3] We would like to thank Dr. Gabriella Vigliocco for suggesting this analysis.

the same clause or in immediately adjacent clauses. In contrast, we assume that the conceptual processes that underlie the production of the first and second mentions are equivalent. While there may be some priming at the conceptual level for the single repeated concept, that priming occurs within the domain of the construction of an entirely new utterance and therefore should not have a discernible effect on the ease of conceptualization. Thus, in utterances like those in (8), the Lexical Level Hypothesis would predict more semantically co-expressive gestures to co-occur with the first mention of *road* then are found for the second mention, since retrieval of the second mention is already facilitated by the first.

(8) a. …and you take the road, the road goes again upwards.
b. …you can go straight ahead or to the left. You go to the left.

An utterance was considered a repetition if the same lexical item was used to refer to the same referent either within the same clause or in the following clause. Repetitions occurred fairly frequently (N=75 pairs); however, only a few had co-expressive gestures on the first or second (or third) mention (N=24). The numbers of first and subsequent mentions of a given lexical item that occurred with a co-expressive gesture are presented in (9) below.

(9) Number of semantically co-expressive gestures produced with the first mention or immediately following mention of the same lexical item with the same referent

	Directions	Destinations	Non-essential	Total
1st mention	4	2	3	9
2nd primed mention	3	5	7	15

As this table shows, there was no tendency to produce a gesture on a first mention rather than on a second or subsequent mention, suggesting that these gestures were not produced to facilitate lexical retrieval.[4] Gesture frequency did not decrease when a lexical item was primed due to repetition. This suggests that these gestures were not produced to facilitate lexical retrieval.

5. Discussion

Most models of speech production agree that conceptualization precedes linguistic formulation. The processes at the conceptual level are varied but include accessing visual, spatial, and encyclopedic information about the world and the ideas to be expressed. If a speaker wants to produce the word *rabbit* she activates many types

[4] For the purposes of this analysis we only investigated the six participants who produced more than one gesture per picture.

of information about the rabbit, including that the fact that it is an animal, has long ears, eats lettuce, etc. These conceptual and semantic features combine to activate a lemma representation corresponding to the word *rabbit*. This lemma representation is associated to morphosyntactic (and possibly also semantic) features of the word.

Given this architecture, there are two ways in which lexical retrieval can be facilitated by gestures. The first route, proposed by the Lexical Level Hypothesis, is a direct activation of the lexical representations by the gesture. According to this hypothesis, gestures prime lexical representations directly, bypassing conceptual level planning. The second possible route is indirect activation of the lexical representation mediated by the conceptual representations, as proposed by the Conceptual Level Hypothesis. According to this proposal, gestures function to activate conceptual level representations and/or aid in conceptual level processes. The representations at the conceptual level then feed activation down to the lexical representation. Given both of these possible accounts for how lexical retrieval might benefit from gestures, the evidence cited in support of the Lexical Level Hypothesis (e.g. that gesture prohibition adversely affects speech production) is naturally explained by both hypotheses.

Thus, to disentangle these two hypotheses of how gestures might facilitate speech production, we attempted to isolate components of our picture description task that were more or less challenging at one of the two processing levels. To investigate the issue of conceptual level facilitation, we contrasted gesture frequency for intrinsic and deictic speakers, arguing that the latter required more computations at the conceptual level then the former. As predicted, we found that the intrinsic speakers produced more gestures than the deictic speakers. Furthermore, we also found that when the demands of the description task were greater, specifically around the decision point, speakers also produced more gestures. Both of these findings support the claims of the Conceptual Level Hypothesis. In contrast, when we compared the number of gestures produced at the first or second mention of a repeated lexical item we found no difference in the gesture frequency. This is counter to the prediction of the Lexical Level Hypothesis.

As suggested earlier, the lexical and conceptual level hypotheses need not be mutually exclusive. It is possible that gesture has a direct facilitative effect at both the lexical and the conceptual levels. The results discussed in the literature are compatible with both hypotheses. Granting this, the results presented in this paper seem to support facilitation only at the conceptual level. However, it should be noted that the results presented here are numerical trends, not statistical significant differences. Furthermore, even finding a statistically significant difference using the correlational design presented above would not prove a causal link between increased processing load and gesture; rather, it would only be suggestive of a relationship at the conceptual level. In addition, it cannot be concluded with certainty that no facilitative effect is present at the lexical level, despite the fact that we found no evidence to support such a conclusion. Most of the lexical items used by the speakers in our study were relatively high frequency and common; they were also repeated over and over again across the sixteen pictures. Thus, it is possible

that this task is not optimally designed to reveal effects of lexical level facilitation. The challenge is to find a task that contrasts lexical properties of words while keeping conceptual properties constant. This is what we attempted to do in the present study. Our results clearly suggest that gestures have a facilitative function at the conceptual level, but whether they may also have a direct, unmediated facilitative function at the lexical level is still unclear despite the absence of evidence to date.

References

Alibali, Martha W., Sotaro Kita, and Amanda J. Young. 2000. Gesture and the Process of Speech Production: We Think, Therefore We Gesture. *Language and Cognitive Processes* 15:593-613.

Butterworth, Brian. 1989. Lexical Access in Speech Production. In William Marslen-Wilson et al. (ed.), *Lexical Representation and Process*. Cambridge, MA: MIT Press.

Butterworth, Brian, and Uri Hadar. 1989. Gesture, Speech and Computational Stages: A Reply to McNeill. *Psychological Review* 96:168-174.

Cohen, Akiba. 1977. The Communicative Functions of Hand Illustrator. *Journal of Communications* 27:54-63.

de Ruiter, Jan-Peter. 1998. Gesture and Speech Production. Ph.D. diss., Max Planck Institute for Psycholinguistics, The Netherlands.

Dell, Gary. 1986. A Spreading Activation Theory of Retrieval in Sentence Production. *Psychological Review* 93:283-321.

Frick-Horbury, Donna, and Robert E. Guttentag. 1998. The Effects of Restricting Hand Gesture Production on Lexical Retrieval and Free Recall. *American Journal of Psychology* 111:43-62.

Kempen, Gerard, and Pieter Huijbers. 1983. The Lexicalization Process in Sentence Production and Naming: Indirect Selection of Words. *Cognition* 14:185-209.

Kendon, Adam. 1994. Do Gestures Communicate?: A Review. *Research on Language and Social Interaction* 27:175-200.

Kita, Sotaro. 2000. How Representational Gestures Help Speaking. In D. McNeill (ed.), *Language and Gesture: Window into Thought and Action*. Cambridge, UK: Cambridge University Press.

Krauss, Robert, Yihsiu Chen, and P. Chawla. 1996. Nonverbal Behavior and Nonverbal Communication: What Do Conversational Hand Gestures Tell Us? *Advances in Experimental Social Psychology* 28:389-450.

Krauss, Robert, Yihsiu Chen, and R. Gottesman. 2000. Lexical Gestures and Lexical Access: A Process Model. In D. McNeill (ed.), *Language and Gesture: Window into Thought and Action*. Cambridge, UK: Cambridge University Press.

Levelt, Willem J. M. 1989. *Speaking: From Intention to Articulation*. Cambridge, MA: MIT Press.
Levelt, Willem J. M., Ardi Roelofs, and Antje S. Meyer. 1999. A Theory of Lexical Access in Speech Production. *Behavioral and Brain Sciences* 22:1-75.
Rauscher, Frances H., Robert Krauss, and Yihsui Chen. 1996. Gesture, Speech and Lexical Access: The Role of Lexical Movement in Speech Production. *Psychological Science* 7:226-231.
Slobin, Dan. 1987. Thinking for Speaking. In J. Aske, N. Beery, L. Michaelis, and H. Filip (eds.), *Proceedings of the Thirteenth Annual Meeting of the Berkeley Linguistics Society*. Berkeley, CA: Berkeley Linguistics Society, 435-444.

FR 4.7 Psycholinguistik
Universität des Saarlandes
Geb. 17.1, Room 1.16
D-66041 Saarbrücken
Germany

melinger@coli.uni-sb.de

Max Planck Institute for Psycholinguistics
Wundtlaan 1
P.O. Box 310
6500 AH Nijmegen
The Netherlands

kita@mpi.nl