Language and Cognition **9** (2017), 1–23. doi:10.1017/langcog.2014.25

Taking simulation semantics out of the laboratory: towards an interactive and multimodal reappraisal of embodied language comprehension*

KASPER KOK
Vrije Universiteit Amsterdam

AND

ALAN CIENKI
Vrije Universiteit Amsterdam and Moscow State Linguistic University

(*Received* 30 *November* 2013 – *Revised* 16 *April* 2014 – *Accepted* 06 *May* 2014 – *First published online* 07 *July* 2014)

ABSTRACT

Recent embodied theories of meaning known as 'simulation semantics' posit that language comprehension engages, or even amounts to, mental simulation. What is meant here by 'language comprehension', however, deviates from the perspectives on interpersonal communication adhered to by researchers in social psychology and interactional linguistics. In this paper, we outline four alternative perspectives on comprehension in spoken interaction, each of which highlights factors that have remained largely outside the current purview of simulation theories. These include perspectives on language comprehension in terms of (i) striving for inter-subjective conformity; (ii) recognition of communicative intentions; (iii) prediction and anticipation in a dynamic environment; and (iv) integration of multimodal cues. By contrasting these views with simulation theories of comprehension, we outline a number of fundamental differences in terms of the kind of process comprehension is assumed to be (passive and event-like versus active and continuous), as well as the kind of stimulus that language is assumed to be (comprising unimodal units versus being multimodal and distributed across conversational turns). Finally, we discuss potential points of connection between simulation semantics and research on

[*] The authors thank Charles Forceville, Benjamin Bergen, and two anonymous reviewers for comments on earlier drafts of this manuscript. They also gratefully acknowledge the Netherlands Scientific Organization (NWO) for a PhDs in the Humanities grant which supported the completion phase of work on this paper.

spoken interaction, and touch on some methodological implications of an interactive and multimodal reappraisal of simulation semantics.

KEYWORDS: simulation semantics, comprehension, interaction, social cognition, multimodality.

1. Introduction

Research on the cognitive basis of language comprehension has developed in considerable ways since the principles of cognitive linguistics were first formulated. Along with furthering the field of 'simulation semantics', researchers have become increasingly concerned with the interactions between meaning and processes in domains such as perception, motor action, and emotion (e.g., Barsalou, 1999; Bergen, 2012; Glenberg & Kaschak, 2002; Zwaan, 2003). The dynamic view of conceptualization as prompted by linguistic input that underlies this paradigm has been explicitly cited as consonant with some views of semantics within cognitive linguistics, such as cognitive grammar (Langacker, 2008: Ch. 14) and conceptual metaphor theory (Gibbs, 2006).

Regrettably, however, the development of this framework has taken place largely independently of two other developments in language research, namely increased interest in (i) the social, contextual, and pragmatic aspects of linguistic communication; and (ii) the variably multimodal nature of spoken language. In this paper, we argue that as a result of divorcing language from its 'canonical encounter', i.e., spoken conversation (Clark, 1973), it is not evident how simulation theories of language comprehension extend to real-life situations. After evaluating the complexity of the process of language comprehension in view of four prevalent perspectives on interpersonal communication, this paper proposes an integrative account that provides the background for assessing the role of mental simulation in everyday language comprehension.

The following presents a brief overview of the state of the art of simulation-based theories of language understanding. Next, we assess how comprehension in a laboratory setting differs from comprehension in spoken interaction and show that the explanatory scope of simulation theories is confined to a non-canonical form of 'language' and an overly simplistic conception of 'comprehension'.[1]

Finally, we discuss a number of connections between simulation semantics and aspects of face-to-face communication, and touch upon some methodological implications.

[1] Many aspects of this critique could just as well be applied to other branches of psycho-linguistics or cognitive linguistics. Simulation semantics is a case in point, chosen because of its claims to go beyond truth-conditional paradigms of meaning in order to account for more real-life use of language.

1.1. SIMULATION THEORIES OF LANGUAGE COMPREHENSION

Simulation-based theories of comprehension have come about in opposition to the premise that meaning relies on abstract, atomic symbols. Copiously supported by experimental research, simulation theories hold that language comprehension engages partial re-enactment of perceptual, motoric, and affective memory traces, insofar as they are relevant to the concept or situation described. Evidence for this proposal has been provided on different levels of linguistic complexity.

1.2. WORD COMPREHENSION

Since the late 1990s, brain imaging studies have demonstrated that processing isolated nouns and verbs recruits substrates of memory systems that correspond to their content. Listening to descriptions of objects, for instance, has been shown to elicit modality-specific neural activation in brain areas related to their perceptual features (Chao, Haxby, & Martin, 1999; Pulvermüller, Mohr, & Schleichert, 1999). Words describing physical actions, likewise, induce activation in motor areas specific to the parts of the body they are performed with (Hauk, Johnsrude, & Pulvermüller, 2004; Hauk & Pulvermüller, 2004; Vigliocco, Warren, Siri, Arciuli, Scott, & Wise, 2006). These and other findings have been taken to support the idea that the 'symbols' on which language operates are abstracted away from embodied experiences, thus grounded in perceptual, motor, and emotive systems (Barsalou, 1999; Evans, 2009).

1.3. SENTENCE COMPREHENSION

Comparable evidence has been obtained in neuroscientific studies where participants were exposed to sentences rather than single words. Desai, Binder, Conant, and Seidenberg (2010), for example, observed modality-specific activation of cortical areas in response to sentences describing motor actions or visual scenes. Tettamanti et al. (2005) and Aziz-Zadeh, Wilson, Rizzolatti, and Iacoboni (2006) report language-induced activation in premotor areas corresponding to foot, hand, or mouth actions, in close correspondence with the type of action described in the stimulus sentences.

More informative results with respect to semantic representation of sentence meaning come from behavioral research. A range of experimental studies has suggested that, in sentence comprehension, people conceptualize perceptual and motor details beyond the propositions presented explicitly. Zwaan, Stanfield, and Yaxley (2002), for example, found that after reading *The ranger saw the eagle in the sky*, participants were faster to respond to an image of a bird with spread wings than to an image of a bird with closed

https://doi.org/10.1017/langcog.2014.25 Published online by Cambridge University Press

wings, whereas this effect was reversed after reading *The ranger saw the eagle in the nest*. Employing analogous research strategies, scholars have provided converging evidence that mental simulations of sentence content encode rigorous perceptual detail of the objects described, including their spatial orientation (Stanfield & Zwaan, 2001), color (Connell, 2007), visibility (Yaxley & Zwaan, 2007), spatial characteristics (Bergen, Lindsay, Matlock, & Narayanan, 2007; Winter & Bergen, 2012), trajectory of motion (Kaschak et al., 2005; Matlock, 2004; Zwaan, Madden, Yaxley, & Aveyard, 2004), and viewpoint (Borghi, Glenberg, & Kaschak, 2004).

Similar interaction effects have been reported in motor and affective domains. Comprehension of sentences describing actions (Klatzky, Pellegrino, McCloskey, & Doherty, 1989) and emotional situations (Havas, Glenberg, & Rinck, 2007) has been found to be sensitive to the comprehender's current bodily state. Conversely, experimental evidence has shown that reading sentences in which physical movement is implied affects subsequent performance of actual motor actions, insofar as the performed and implied actions are compatible (Bergen & Wheeler, 2005; Glenberg & Kaschak, 2002; Zwaan & Taylor, 2006). The motor simulations involved in processing action descriptions, furthermore, appear to be sensitive to the AFFORDANCES of the objects described, i.e., the way they can be manipulated or interacted with (Borghi & Riggio, 2009; Chwilla, Kolk, & Vissers, 2007; Kaschak & Glenberg, 2000; Masson, Bub, & Warren, 2008).

In an attempt to bring together these and other findings, Zwaan (2003) hypothesizes that during sentence comprehension, people construct vicarious and holistic mental simulations of the expressed content in an IMMERSED fashion, i.e., as if they are actually part of the referential scene. Similar hypotheses have been set forward by Barsalou (2003, 2005).

1.4. DISCOURSE COMPREHENSION

The study of language comprehension on the level of stretches of written discourse longer than single sentences has developed largely independently of the research mentioned so far. Discourse comprehension research has focused mainly on people's ability to keep track of the local and global coherence in a text and the kinds of inferences they draw in doing so. Such inferences involve the spatial, temporal, and causal structure of the described situations, as well as the traits, goals, and motives of the relevant characters (Zwaan, Magliano, & Graesser, 1995). The mental constructs subsuming the totality of these inferences, in addition to the information explicitly presented, are often referred to as SITUATION MODELS (Graesser, Millis, & Zwaan, 1997; Van Dijk & Kintsch, 1983; Zwaan & Radvansky, 1998). The neurocognitive underpinnings of situation models have only relatively

https://doi.org/10.1017/langcog.2014.25 Published online by Cambridge University Press

recently been gaining serious attention, along with advances in research on embodied sentence processing. Zwaan (2009, p. 1145), for instance, argues that simulation theories of comprehension "fill the theoretical gap" between situation model theories of event relations and embodied theories of conceptualization. Empirical support for recruitment of sensorimotor systems in discourse comprehension is available (e.g., Speer, Zacks, & Reynolds, 2007; Wallentin, Nielsen, Vuust, Dohn, Roepstorff, & Lund, 2011), but the relation between simulation semantics and situation model theory currently remains underspecified (see also Section 3.2).

1.5. LANGUAGE COMPREHENSION?

Simulation semantics, in its current state, is not devoid of controversy. Current debates mostly center around the questions whether abstract language engages mental simulation in the same way as concrete language (e.g., Dove, 2010) and whether simulation actually has a FUNCTIONAL role with respect to comprehension (Mahon & Caramazza, 2008; Willems & Francken, 2012). A question less often addressed is what exactly it means to claim that simulation semantics provides 'a theory of language comprehension'. To what degree, if at all, is the way people process isolated, written excerpts of language in the laboratory similar to the way people comprehend each other in everyday life? Some researchers of (narrative) discourse comprehension have argued that the stimuli they deploy provide a good model for everyday language use, claiming that:

> [n]arrative text has a close correspondence to everyday experiences in contextually specific situations, […] both narrative texts and everyday experiences involve people performing actions in pursuit of goals, the occurrence of obstacles to goals, and emotional reactions to events. (Graesser, Singer, & Trabasso, 1994, p. 372)

Whereas it might be true that narrative text comprehension, given its contextualized character, is more natural than the comprehension of isolated sentences or words, it is still, in many ways, an artificial form of language, encountered in relatively late stages of language acquisition. The ecological validity of the research discussed in the previous sections, therefore, might not be evident as previously supposed. In line with Willems and Francken (2012), we contend that the role of mental simulation in the process of language comprehension is to be understood against the background of a situated and interactive view on language use. As a first step toward laying the foundations of such a theory, the following sections review four different perspectives on language comprehension in everyday spoken interaction.

https://doi.org/10.1017/langcog.2014.25 Published online by Cambridge University Press

2. Comprehension in the laboratory vs. comprehension in interpersonal communication

Social-interactional factors have long been considered beyond, or at most peripheral to, cognitive linguistic approaches to meaning (see Geeraerts, 2010). Although critiques of the cognitivist take on language processing are not new (cf. Parker, 1992), the undisputable importance of such factors for the study of linguistic meaning has only recently become more widely accepted. As Croft (2009, p. 395) argues:

> [T]he foundations of cognitive linguistics […] are too solipsistic, that is, too much 'inside the head'. In order to be successful, cognitive linguistics must go 'outside the head' and incorporate a social-interactional perspective on the nature of language.

Literature from social psychology, indeed, demonstrates that language comprehension in interpersonal communication encompasses much more than text comprehension does, and can be approached from a variety of perspectives (Krauss & Fussell, 1996). In addition to constructing semantic representations, language comprehension in face-to-face interaction draws upon various pragmatic and social-interactive capacities. Moreover, it involves the dynamic integration of verbal, intonational, and gestural cues. In the following, we outline four perspectives on how 'language comprehension' in interpersonal communication can be defined, highlighting factors that have remained largely outside the current purview of simulation theories of comprehension.

2.1. COMPREHENSION AS INTERSUBJECTIVE CONFORMITY

Understanding a sentence like *The ranger saw the eagle in the sky* in a decontextualized,[2] experimental setting entails taking hold of the TYPE of situation this sentence may refer to. In interpersonal communication, by contrast, comprehension involves pursuing INTERSUBJECTIVE CONFORMITY; that is, comprehenders need to assess which particular TOKENS of entities or events the speaker is most likely to refer to given the current context (e.g., a particular eagle or a particular instance of seeing that eagle). The human capacity for reference resolution in contextualized communicative situations, thus, is integral to situated language comprehension.

[2] Of course a laboratory is a context in itself, a special one that can make readers more aware of a text's potential meanings than they normally might be, but what is referred to here is the lack of much co-text or a context in which the sentence would make sense in the immediate environment.

https://doi.org/10.1017/langcog.2014.25 Published online by Cambridge University Press

One of the first accounts to take the relation between context and referential meaning seriously was Barwise and Perry's (1981, 1983) influential SITUATION SEMANTICS, which proposed that the meaning of an utterance is determined by the set of alternative interpretations that the communicative situation avails, and that the correct interpretation can be deduced by means of parameter setting. Somewhat independently, reference disambiguation has become a topic of interest in social psychological approaches to language, as an aspect of 'grounding' (Clark & Brennan, 1991; Clark & Marshall, 1981). On Clark's (1996 and elsewhere) account, language users understand referential expressions through assessment of their COMMON GROUND, i.e., the degree to which current beliefs, knowledge, and suppositions are shared. Reference resolution, on this view, can be thought of as an active search process, whereby the search space is restricted to the assumed overlap in viewpoint and experiential memory between the interlocutors. Others have hypothesized that much of the problem of reference resolution may be resolved in a more automatic way, by dint of ad-hoc referential routines established during dialogue (Pickering & Garrod, 2004).

2.2. COMPREHENSION AS INTENTION RECOGNITION

Plausibly reminiscent of the computational paradigm of the mind dominant during the previous decades, many experimental studies on language processing implicitly assume language's main goal to be 'the transfer of information' (cf. Shannon & Weaver, 1948). As Austin (1962) and other pragmaticists have pointed out, however, one can do much more with words than assert information. A more contemporary view, as proposed in Relevance Theory (and elsewhere), is that linguistic expression serves to bring about "contextual effects in an individual" (Sperber & Wilson, 1986, p. 265). The primary function of linguistic utterance, accordingly, is to elicit either an overt (e.g., verbal) response by the addressee or a covert effect (e.g., a change of beliefs or intentions).

From the addressee's point of view, this 'intentionalist' paradigm of communication implies that the essence of language understanding is not to seize the literal (referential) meaning of an utterance, but to successfully infer what effect the speaker aims to bring about. Akin to the debate on the cognitive underpinnings of common ground assessment, theories on the mechanisms underlying intention recognition vary along the automatic–effortful continuum (Carruthers & Smith, 1996). Some 'categorical' aspects of intention are commonly believed to be understood directly through routinized associations with linguistic elements (e.g., through grammatical marking of illocutionary force), while more ad-hoc aspects of the speaker's communicative intention have been argued to recruit effortful computations,

https://doi.org/10.1017/langcog.2014.25 Published online by Cambridge University Press

taking into account presumptions about the speaker's current mental state (Goldman, 1992).

2.3. COMPREHENSION AS DYNAMIC ANTICIPATION

A third limitation of many current accounts of comprehension is their conception of language use as a unidirectional event. Despite abundant criticism on the 'myth of the isolated mind' inherent in such approaches (Bakhtin, 1981; Clark, 1973, 1996; Linell, 2007; Stolorow & Atwood, 1992), it is still common practice in psycholinguistic and cognitive linguistic research to consider language processing as taking place inside one individual brain. Alternative, DIALOGICAL approaches to comprehension, by contrast, hold that linguistic communication should not be regarded as an encapsulated event, but as a dynamic and interactive process in which meanings and intentions are cooperatively negotiated. Linguistic communication, according to this view, can be characterized as a JOINT ACTIVITY: a process whereby interlocutors collaborate to achieve a shared conception of the ongoing discourse.

Within dialogical models, a discrepancy exists in terms of the types of interaction considered relevant to the dialogue. A first type of model is predominantly concerned with interactions among interlocutors themselves, whereby dialogue is regarded as a process of jointly constructing shared discourse models (Bakhtin, 1981; Garrod & Anderson, 1987; Garrod & Pickering, 2004). A second, more radical version of 'dialogism' maintains that an accurate model of language use should not only take interactions between interlocutors into account, but instead take the dynamics of the communicative setting in which the interaction takes place as a starting point (e.g., Linell, 1998, 2007). Language comprehension and other cognitive capacities, on this view, are to be understood as residing in a continual dialectic between the interlocutors and their environment, and to be studied in relation to the situational continuum in which they are embedded: "Rather than being the seat of epistemically private mental representations, the brain functions to regulate the body's interactions with its ecosocial environment" (Thibault, 2005, p. 152). Neither the semantic nor the pragmatic–intentional dimensions of comprehension, accordingly, should be seen as encapsulated events. Rather, these aspects of comprehension can be thought of as instrumental to the more general cognitive capacities for projection and anticipation in dynamic interaction. Or, as Linell (2007, p. 611) puts it:

> [Meaning] potentiality is related to creativity and adaptability, to the principled capacity of language to meet the communicative needs of ever changing situations […] To understand an utterance in real time, we must

https://doi.org/10.1017/langcog.2014.25 Published online by Cambridge University Press

be able to predict the continuation and project one's own and others' possible next actions.

2.4. COMPREHENSION AS MULTIMODAL

A final dimension of face-to-face language comprehension which has largely remained on the background pertains to language's multimodal nature. A few examples aside (e.g., Richardson, Spivey, Barsalou, & McRae, 2003; Winter & Bergen, 2012), cues for constructing mental simulations in the experiments discussed all have the form of written text. The foundations of simulation theories, thus, rely on a form of language that is SENTENCE-BASED and UNIMODAL, in sharp contrast to the form language takes in spoken interaction (cf. Chafe, 1994: Ch. 2; Linell, 1982/2005; Ochs, Schegloff, & Thompson, 1996).

A first difference is that extemporaneous spoken speech, being a time-constrained activity, involves an on-line process of 'information packaging' (Chafe & Tannen, 1987). Chafe (1994, p. 109) and others have posited that speakers distribute their messages over prosodically delineated INTONATION UNITS so as to conform to their own processing constraints and those of the addressee. Using the term 'tone unit' (TU) instead of intonation unit, Altenberg (1987, p. 46) claims that "it is in terms of TUs — rather than any specific grammatical unit — that speakers organize and present information in discourse, and it is through TUs that listeners perceive and understand this information".

Information packaging has profound implications for discourse-anticipatory dimensions of language comprehension. Prosodic contours are known to mark an utterance's information structure and to be revealing with respect to the flow and continuation of the discourse (Bolinger, 1986; Brazil, 1997). Listeners, as demonstrated experimentally (e.g., Swerts & Geluykens, 1994), exploit melodic and pausal cues to process local and global aspects of discourse structure. Comprehension of spoken language, thus, is to be seen as an incremental process which imposes different processing constraints than written language.

Second, face-to-face communication allows for extensive use of ostensive behaviors other than oral expression, including manual and facial gestures. The type of gestures most often studied in this context, those with a representational function, primarily pertains to semantic dimensions of comprehension (e.g., Beattie & Shovelton, 1999; McNeill, 1992). Such gestures can be communicative in various ways (Kendon, 2004), for instance by providing spatial content in a way that speech is not suited for, by providing information that is additional to what is conveyed verbally, or by providing

https://doi.org/10.1017/langcog.2014.25 Published online by Cambridge University Press

additional cues in case speech comprehension is difficult (Hostetter, 2011; Kendon, 1994).

In addition to their relevance to semantic processes, co-speech gestures have been associated with a range of functions in the realm of pragmatics. Kendon (2004, pp. 158–159) mentions three main pragmatic functions of gesture, namely modal functions (which alter the frame in terms of which what is being said is to be interpreted), performative functions (that indicate the kind of speech act the person is engaged in), and parsing functions (e.g., marking the logical structure of what is being uttered). In the latter function, speakers may, for example, use their hands to contrast two opposing positions in a debate, or to sum up a list of points. Experimental studies, furthermore, have foregrounded gesture's role in disambiguation of verbal expressions and interactive grounding (Clark & Krych, 2004; Holler & Beattie, 2003; Kelly, Özyürek, & Maris, 2010). In interactional terms, such gestures allow the addressee to predict the upcoming material and prepare their own reactions.

Altogether, intonation and co-speech gestures are (often simultaneously) relevant to comprehension in terms of semantic, pragmatic, and anticipatory aspects of spoken language comprehension.

2.5. A RECONCILED VIEW

The previous sections have set forward two fundamental differences between language comprehension inside and outside the laboratory. The first concerns the KIND OF PROCESS that comprehension is (active and continuous rather than passive and event-like). The second concerns the KIND OF STIMULUS that language is (multimodal to varying degrees and distributed across conversational turns, rather than consisting of unimodal units). Language comprehension in face-to-face communication, all in all, is a dynamic and multifaceted activity, accomplished on the basis of verbal as well co-verbal signaling.

How are these facts to be reconciled into a coherent processing model? Notwithstanding that there exists some tension between the perspectives outlined in the sections above in terms of their underlying presumptions, they are not fully incompatible. Rather, they can be seen as mirroring different layers of a hierarchically organized processing architecture. In line with Clark's (1999) view of speech acts as comprising 'action ladders', the multilayeredness of comprehension can be thought of as reflecting language's role as a COORDINATION DEVICE for communication, and communication's role as a coordination device for other types of joint action (cf. Clark, 1996; Croft, 2009). Comprehension in the sense of establishing intersubjective conformity, accordingly, is at least to some extent conditional on intention recognition, which can be considered subordinate to the more general skill of engaging in dynamic interaction.

https://doi.org/10.1017/langcog.2014.25 Published online by Cambridge University Press

These different types of process, however, are not to be regarded as modular or merely sequentially organized. Ample evidence has shown that linguistic and more general (socio-)cognitive processes continually interact in a top-down fashion as well (e.g., Hagoort, Hald, Bastiaansen, & Petersson, 2004; Van Berkum, Van Den Brink, Tesink, Kos, & Hagoort, 2008). A felicitous processing model should therefore acknowledge that comprehension emerges out of the bi-directional interplay between linguistic and more general communicative capacities, and, from a mechanistic point of view, is to be defined "across multiple coupled dynamical systems" (Wilson & Golonka, 2013, p. 10).

Figure 1 provides a simplistic sketch of a model that incorporates these considerations: general socio-cognitive processes and more specific linguistic processes are portrayed as constituting a hierarchically organized, mutually interactive network.[3] The term 'analysis' is used after Bergen and Chang (2005), referring to the process of extracting the parameters according to which a mental simulation is performed from the perceived utterance. The other three levels correspond to the dimensions of comprehension discussed in this section. The arrows on the left indicate that all different subcomponents of face-to-face language comprehension are potentially sensitive to gesture, intonation, and other co-verbal behaviors.

This model is contrasted with the 'experimental' take on language comprehension, where co-verbal, discursive, and social-contextual factors are typically factored out. This comparison, notably, is not meant to suggest that the experimenters in question do not (at least in theory) acknowledge the importance of social-pragmatic processes, but rather to point out that factoring out such facets of comprehension at the benefit of experimental control may come at the expense of the ecological validity of this type of research.

Does this mean, then, that social-pragmatic processes and co-verbal aspects of communication are to be a concern for (simulation) semanticists? Many would contend that the inclusion of such factors blurs the scope of what semantics is about. For two related reasons, however, these issues are worth considering. The first is that, from a cognitive point of view, the scope of semantics is barely delineable in the first place. As argued by Langacker (1987, 1997) and others, the meaning of contextualized utterances is never devoid of pragmatics. Because the usage events from which linguistic meanings are abstracted have predominantly taken place in social-interactive settings,

[3] The modular organization of this model is merely an illustrative artifact, and is not meant to imply that these steps sensibly be regarded as encapsulated mechanisms.

https://doi.org/10.1017/langcog.2014.25 Published online by Cambridge University Press

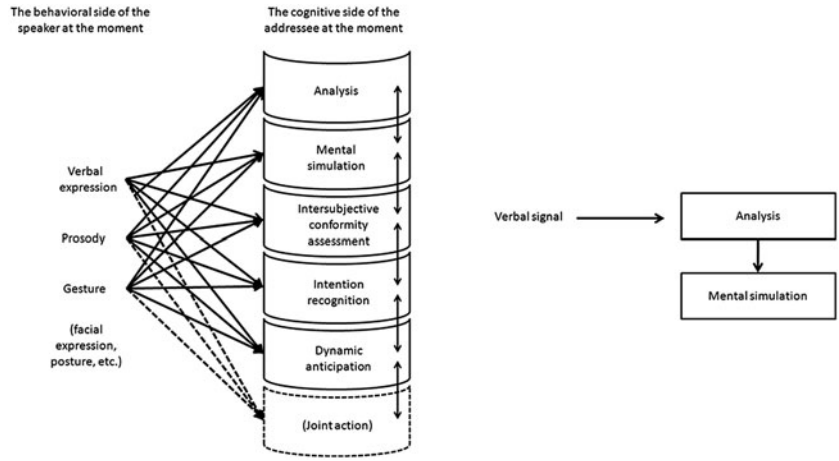

Fig. 1. A schematic comparison of the complexity of language comprehension (at a given moment) outside (left) and inside (right) the laboratory. Notably, due to representational limits, this model does not accurately capture the dynamic and distributed nature of comprehension in dialogue.

all utterances have an inherent socio-pragmatic import. A second reason is that simulation theories are often not only presented as an approach to semantics, but also explicitly put forward as a theory of language comprehension (e.g., Bergen & Chang, 2005; Glenberg & Robertson, 1999; Zwaan, 2003). If this ambition is taken seriously, the interfaces between semantic, social-pragmatic, and multimodal aspects of communication need more thorough examination. The following section raises a number of issues and questions relevant to pursuing this goal.

3. Simulation semantics in the context of interactive, spoken language comprehension

Here we discuss some issues and open questions that are relevant to assessing the explanatory scope of current simulation theories in the light of a broader, interactional perspective on comprehension.

3.1. MENTAL SIMULATION AND MENTALIZING

By factoring out the communicator and situational context, experimental studies such as those reviewed earlier limit their participants' freedom for meaning construction to the individual mind. This arguably makes an unnatural appeal to the participant's language resources, which have been claimed to be "designed to be completed only in situated meaning-making"

https://doi.org/10.1017/langcog.2014.25 Published online by Cambridge University Press

(Linell, 2007, p. 611). In other words, this research can be said to capture the subjective, but not the intersubjective nature of meaning.

As discussed before, the human capacity for social-pragmatic dimensions of comprehension, such as the inference of others' viewpoint and communicative intentions, has been ascribed to automatic, associative systems as well as more effortful 'mentalizing' mechanisms (cf. Carruthers & Smith, 1996). We will here discuss the relation of these types of mechanism to language-driven mental simulation in turn.

Associative accounts of intention recognition maintain that mental states are in essence inaccessible, and that social behavior is essentially understood by virtue of regularities (or 'rules') in social-interactive experiences (Gopnik, 1995). This type of account is quite readily commensurable with the principles of simulation semantics, insofar as mental simulations are not approached as static, internal entities, but as "dynamic, generalized associations which always act relative to the environment" (Robinson, 2000, p. 260). That is, simulation semantics has the potential to extend to socially embedded language use, simply by acknowledging that the experiential knowledge people re-enact during language comprehension does not only involve perception and bodily action, but also contextualized patterns of goal-oriented communication with other individuals. This extension, in fact, follows quite naturally from the usage-based paradigm that lies at its core.

Other theorists have argued that intention recognition amounts to generating a MENTAL SIMULATION of the interlocutor's behavior, as to "replicate, mimic, or impersonate with the mental life of the target agent" (Gallese & Goldman, 1998, p. 497). The question of how this relates to language-driven simulation, inevitably raised by this terminological resemblance, is subject of controversy. Some have argued that both types of 'simulation' draw on embodied mechanisms and ultimately reside in patterns of activity in substrates of the mirror neuron system (e.g., Gallese, 2007). Evidence from neuroimaging (Willems, de Boer, de Ruiter, Noordzij, Hagoort, & Toni, 2010), electrophysiology (Egorova, Shtyrov, & Pulvermüller, 2013), and research on aphasia (Willems, Benn, Hagoort, Toni, & Varley, 2011), however, suggests that semantic and pragmatic aspects of comprehension rely on largely distinct neural substrates (for a review, see Willems & Varley, 2010). The terminological resemblance of the two 'simulation theories', hence, does not seem to reflect full functional or neural overlap.

Various questions still remain with respect to the relation between simulation semantics and social-cognitive capacities. A first issue concerns the notion of 'inference': do the different types of inference involved in sentence comprehension (Section 1.3) — situation model construction (Section 1.4) and intention recognition (Section 2.2) — rely on the same (type of) neural mechanism? A second question concerns the notion of 'perspective',

https://doi.org/10.1017/langcog.2014.25 Published online by Cambridge University Press

which seems to play an important role in the construction of mental simulations (Borghi, et al., 2004; Brunyé, Ditman, Mahoney, Augustyn, & Taylor, 2009; Zwaan, 2003), as well as in social aspects of communication (Krauss & Fussell, 1988; Sperber & Wilson, 1986): To what extent does the human capacity for perspective reallocation constitute an interface between mental imagery and mentalizing capacities? Finally, one may wonder to what extent the vicarious, immersed, character of mental simulations has a functional role with respect to pragmatic inference. Do the representational details of language-driven mental simulations constitute a source for the hearer to draw upon in understanding the behavioral implications of an utterance, or is the vivacity of language-induced imagery merely epiphenomenal to the brain's associative nature, and irrelevant for social aspects of comprehension?

3.2. MENTAL SIMULATION AND DIALOGISM

Some dialogue-oriented theories have proposed the ultimate goal of communication to be the ALIGNMENT OF SITUATION MODELS (Menenti, Pickering, & Garrod, 2012; Pickering & Garrod, 2004), and have argued that embodiment (of various kinds) plays a substantial role in accomplishing this (Pickering & Garrod, 2009). This view of dialogical language understanding, as the co-construction of shared representations, is as of yet far from commensurable with simulation semantics. Little is known about the role of mental simulation in the process of integrating individual utterance meanings into a broader representation of the ongoing dialogue (despite Zwaan's, 2009, p. 1145, optimism that simulation theories have the potential to 'bridge the gap' between sentence comprehension and discourse comprehension). For instance, we do not know much about how long mental simulations persist over time as the dialogue unfolds, and what memory systems modulate their accessibility during turn-taking. In addition, there is a paucity of research on the way mental simulations relate to aspects of dialogue such as back-channel responses, ellipses, and interactive repair strategies.

In view of the discussion of 'strong dialogism' in Section 2.3, moreover, one might argue that dialogical language comprehension is best approached without recourse to representational notions such as 'situation model' in the first place. As Linell (2007) argues, when acknowledging that language understanding is part of a continual series of interactions with the environment:

> the emphasis shifts from REPRESENTATION TO CONTROL, INTERACTION AND INTERVENTION. While we surely need knowledge of and assumptions about the world, the various corresponding 'representations' are largely subordinated to interaction and intervention in the world. (Linell, 2007, p. 613, emphasis in the original)

https://doi.org/10.1017/langcog.2014.25 Published online by Cambridge University Press

Mental simulations and situation models, accordingly, may have no reality independent of the predictions that they allow the language user to make and the actions that they serve to prepare. This shifts the burden (for the comprehender) from internal cognitive processes to the direct employment of perceptual and actional resources for engaging in (joint) action (Varela, Thompson, & Rosch, 1991; Wilson & Golonka, 2013). Comprehension, accordingly, is not to be seen as "the calculations and representation of a knowledge structure in the mind", but rather as "the state of the cognitive system at a certain point in time in relation to the world around it" (Robinson, 2000, p. 260).

Such models ostensibly align with the view that mental simulations are instrumental to action preparation and prediction (e.g., Barsalou, 2009; Willems & Hagoort, 2007; Zwaan & Taylor, 2006). However, important theoretical issues have to be resolved for these perspectives to be truly commensurable. Because stronger versions of dialogism (e.g., Varela et al., 1991; Wilson & Golonka, 2013) reject the notion of mental representation *tout court*, they are in essence at odds with the basic principles of simulation semantics. Whether this tension can be resolved in a constructive fashion remains a topic of dispute, to which some have expressed optimism. Van Elk, Slors, and Bekkering (2010), for instance, propose a PROCEDURAL rather than representational interpretation of mental simulation theory. Rączaszek-Leonardi (2009), alternatively, proposes a view of linguistic symbols as CONSTRAINTS on interactional dynamics. These views invite reframing the debate on embodiment in terms of whether grounded (embodied) symbols constrain interaction differently from abstract (disembodied) symbols.

3.3. MENTAL SIMULATION, MULTIMODALITY AND APPREHENSIVE FLEXIBILITY

The potential connections between simulation semantics and gesture research are plentiful, from the point of view of both language production and comprehension (Marghetis & Bergen, in press). Hostetter and Alibali's (2008) influential GESTURE AS SIMULATED ACTION FRAMEWORK, for instance, hypothesizes that co-speech gestures originate in spatial-motoric simulations performed during speech production. Conversely, it has been argued that, during comprehension, verbal and co-verbal components of expression are integrated into a shared representation of meaning through continual interactions (Kelly et al., 2010; Özyürek, Willems, Kita, & Hagoort, 2007). In view of the apparent involvement of mirror neurons in gesture perception (Bernardis & Gentilucci, 2006; Montgomery, Isenberg, & Haxby, 2007; Skipper, Goldin-Meadow, Nusbaum, & Small, 2007) it has furthermore been hypothesized that the representational resources involved in verbally

https://doi.org/10.1017/langcog.2014.25 Published online by Cambridge University Press

evoked motor imagery can be 'merged' with the motor resonance elicited by gestures: "activation from the gesture can summate with activation from speech and contextual information to substantially reduce uncertainty as to what needs to be simulated" (Glenberg & Gallese, 2012, p. 917).

Interestingly, the cortical networks involved in understanding manual behaviors have been proven to be sensitive to contextual factors such as the cultural background of the speaker (Molnar-Szakacs, Wu, Robles, & Iacoboni, 2007) and the communicative relevance of these behaviors (Skipper et al., 2007). Skipper et al. (2007, p. 274), in discussing these findings, speculate that:

> if the behavioral goal involves understanding a sentence when speech-associated gestures can be observed, then areas of the cortex involved in the execution of hand movements and semantic aspects of language comprehension are likely to constitute the mirror system [...] the human mirror system dynamically changes according to the observed action, and the relevance of that action, to understanding a given behavior.

This proposed flexibility of resources for comprehension is in line with Cienki's (2012) hypothesis that the scope of behaviors taken into account by a language user depends on the behaviors' relevance to the current situation. In other words, people variably employ various audible and/or visible behaviors for communicative aims depending on cognitive and contextual affordances and constraints (these constitute what Cienki terms the producer's SCOPE OF RELEVANT BEHAVIORS). The hypothesis continues that, likewise, those attending to speakers apprehend a variable scope of the producer's audible and/or visible behaviors as relevant for communication, varying sometimes moment by moment. Applied to language comprehension, this renders the prediction that different cues for comprehension (e.g., elements of speech, gestures, and prosody) evoke sensorimotor simulations only to the extent that these contribute to engagement in the ongoing interaction. This can offer an explanation to the finding that the semantic resources deployed in comprehension are highly sensitive to task demands and various forms of context (e.g., Sato, Mengarelli, Riggio, Gallese, & Buccino, 2008; Van Dam, Rueschemeyer, Lindemann, & Bekkering, 2010).

3.4. TOWARDS METHODOLOGICAL CONVERGENCE

A fully ecological reassessment of simulation theories, in all directions proposed simultaneously, may not be a realistic objective. The issues outlined in this paper, nonetheless, can inspire future researchers to move in the direction of theorizing about and studying a more natural form of language, as well as a more interactive model of comprehension. The first may simply involve extensions of current experimental studies. By supplementing stimuli

https://doi.org/10.1017/langcog.2014.25 Published online by Cambridge University Press

with a co-verbal (gestural or prosodic) dimension, a better understanding can be gained of the relation between sensorimotor processes and the variably multimodal nature of spoken language.

Arriving at a more interactive notion of comprehension is a more substantial challenge. A first step is to take context and task demands more seriously. In order to avoid the caveat of seeing comprehension as a process that takes place entirely inside an individual's brain, a research program akin to that outlined by Wilson and Golonka (2013) can be of help. This program dictates a careful analysis of experimental task demands and the various resources that may be relevant for satisfying them. In terms of comprehension research, this entails a closer inquiry of the availability of situated and long-term memory resources in during (experimental) comprehension tasks. The taxonomy of different 'levels of situational embedding' proposed by Zwaan (2014) can be a starting point for analyzing (or modulating) the availability of such resources.

A third direction is to approach comprehension as an activity, rather than a passive process. This involves more than just augmenting current experimental paradigms. Rather, there is a need for future studies to incorporate interactive settings, where participants engage in a shared activity. Inferences on the nature and role of the semantic resources recruited in such interactions may unavoidably be more indirect than those in more traditional lab-based tasks (e.g., to be derived from eye-tracking or gesture analysis), but can be an important diagnostic of the ecological value of previously obtained results. Modeling the way (dis)embodied concepts constrain interactional dynamics, e.g., from a dynamical systems point of view (cf. Dale, Fusaroli, Duran, & Richardson, 2014; Rączaszek-Leonardi, 2009), can furthermore initiate a better understanding of the connection between representation-based and dynamical accounts of comprehension, as well as the role that sensorimotor grounding plays in this respect.

4. Conclusion

Clark (1997, p. 594) asserts that:

> [l]anguage understanding is so complex that we have had to cut it into model-sized pieces to study it. But in cutting it up we have also made a number of idealizations, and many of these have become dogmas – premises we take as gospel.

The issues discussed in this paper are natural consequences of this development: whereas taking language comprehension into the laboratory has given rise to many insights on the cognitive nature of semantics, it has at the same time removed language from its natural form and environment.

https://doi.org/10.1017/langcog.2014.25 Published online by Cambridge University Press

As a consequence, current theories of language comprehension are based on impoverished notions of 'language' and 'comprehension'. In this paper, we have argued that the external validity of experimentally based accounts of language comprehension is more questionable than generally supposed: it is by no means evident whether and how experimentally obtained results on the involvement of mental simulation in comprehension extend to real-life situations, where communication is multimodal to varying degrees and embedded in an interactional setting. Potential connections between simulation semantics and social-pragmatic aspects of comprehension have been discussed, but need much more examination. New types of experimental stimuli and more interactive research paradigms are needed in order to better understand the role of mental simulation in everyday face-to-face language comprehension. Most importantly, language needs to be studied as a variably multimodal phenomenon and the research needs to reflect its primary status as a vehicle for communication in a dynamic environment.

REFERENCES

Altenberg, B. (1987). *Prosodic patterns in spoken English: studies in the correlation between prosody and grammar for text-to-speech conversion*. Lund: Lund University Press.

Austin, J. (1962). *How to do things with words*. London: Oxford University Press.

Aziz-Zadeh, L., Wilson, S., Rizzolatti, G., & Iacoboni, M. (2006). Congruent embodied representations for visually presented actions and linguistic phrases describing actions. *Current Biology*, **16**(18), 1818–1823.

Bakhtin, M. (1981). *The dialogic imagination*. Austin: University of Texas.

Barsalou, L. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, **22**(4), 577–609.

Barsalou, L. (2003). Situated simulation in the human conceptual system. *Language and Cognitive Processes*, **18**(5/6), 513–562.

Barsalou, L. (2005). Situated conceptualization. In H. Cohen & C. Lefebvre (Eds.), *Handbook of categorization in cognitive science* (pp. 619–650). St Louis: Elsevier.

Barsalou, L. (2009). Simulation, situated conceptualization, and prediction. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **364**, 1281–1289.

Barwise, J., & Perry, J. (1981). Situations and attitudes. *Journal of Philosophy*, **78**(11), 668–691.

Barwise, J., & Perry, J. (1983). *Situations and attitudes*. Cambridge, MA: MIT Press.

Beattie, G., & Shovelton, H. (1999). Mapping the range of information contained in the iconic hand gestures that accompany spontaneous speech. *Journal of Language and Social Psychology*, **18**(4), 438–462.

Bergen, B. K. (2012). *Louder than words: the new science of how the mind makes meaning*. New York: Basic Books.

Bergen, B. K., & Chang. (2005). Embodied Construction Grammar in simulation-based language understanding. In J-O Östman & M Fried (Eds.), *Construction Grammars: cognitive grounding and theoretical extensions* (pp. 147–190). Amsterdam & Philadelphia: John Benjamins.

Bergen, B. K., Lindsay, S., Matlock, T., & Narayanan, S. (2007). Spatial and linguistic aspects of visual imagery in sentence comprehension. *Cognitive Science*, **31**(5), 733–764.

Bergen, B. K., & Wheeler, K. (2005). Sentence understanding engages motor processes. Paper presented at the Proceedings of the 27th Annual Conference of the Cognitive Science Society, Mahwah, NJ.

https://doi.org/10.1017/langcog.2014.25 Published online by Cambridge University Press

Bernardis, P., & Gentilucci, M. (2006). Speech and gesture share the same communication system. *Neuropsychologia*, **44**(2), 178–190.
Bolinger, D. (1986). *Intonation and its parts: melody in spoken English*. Stanford: Stanford University Press.
Borghi, A., Glenberg, A., & Kaschak, M. P. (2004). Putting words in perspective. *Memory & Cognition*, **32**(6), 863–873.
Borghi, A., & Riggio, L. (2009). Sentence comprehension and simulation of object temporary, canonical and stable affordances. *Brain Research*, **1253**, 117–128.
Brazil, D. (1997). *The communicative value of intonation in English*. Cambridge: Cambridge University Press.
Brunyé, T., Ditman, T., Mahoney, C., Augustyn, J., & Taylor, H. (2009). When you and I share perspectives pronouns modulate perspective taking during narrative comprehension. *Psychological Science*, **20**(1), 27–32.
Carruthers, P., & Smith, P. (1996). *Theories of theories of mind*. Cambridge: Cambridge University Press.
Chafe, W. (1994). *Discourse, consciousness, and time: the flow and displacement of conscious experience in speaking and writing*. Chicago: University of Chicago Press.
Chafe, W., & Tannen, D. (1987). The relation between written and spoken language. *Annual Review of Anthropology*, **16**, 383–407.
Chao, L., Haxby, J., & Martin, A. (1999). Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. *Nature Neuroscience*, **2**(10), 913–919.
Chwilla, D., Kolk, H., & Vissers, C. (2007). Immediate integration of novel meanings: N400 support for an embodied view of language comprehension. *Brain Research*, **1183**, 109–123.
Cienki, A. (2012). Usage events of spoken language and the symbolic units we (may) abstract from them. In J. Badio & K. Kosecki (Eds.), *Cognitive processes in language* (pp. 149–158). Bern: Peter Lang.
Clark, H. (1973). *Space, time, semantics, and the child*. New York: Academic Press.
Clark, H. (1996). *Using language*. Cambridge: Cambridge University Press.
Clark, H. (1997). Dogmas of understanding. *Discourse Processes*, **23**(3), 567–598.
Clark, H. (1999). On the origins of conversation. *Verbum*, **2**, 147–161.
Clark, H., & Brennan, S. (1991). Grounding in communication. In L. B. Resnick, J. M. Levine, & S. D. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 127–149). Washington, DC: APA Books.
Clark, H., & Krych, M. (2004). Speaking while monitoring addressees for understanding. *Journal of Memory and Language*, **50**(1), 62–81.
Clark, H., & Marshall, C. (1981). Definite reference and mutual knowledge. In A. K. Joshi, B. L. Webber, & I. A. Sag (Eds.), *Elements of discourse understanding* (pp. 10–63). Cambridge: Cambridge University Press.
Connell, L. (2007). Representing object colour in language comprehension. *Cognition*, **102**(3), 476–485.
Croft, W. (2009). Towards a social cognitive linguistics. In V. Evans & S. Pourcel (Eds.), *New directions in cognitive linguistics* (pp. 395–420). Amsterdam: John Benjamins.
Dale, R., Fusaroli, R., Duran, N., & Richardson, D. (2013). The self-organization of human interaction. In H. Ross (Ed.), *Psychology of Learning and Motivation, Vol. 59* (pp. 43–95). Online: <http://www.sciencedirect.com/science/article/pii/B9780124071872000022>.
Desai, R., Binder, J., Conant, L., & Seidenberg, M. (2010). Activation of sensory–motor areas in sentence comprehension. *Cerebral Cortex*, **20**(2), 468–478.
Dove, G. (2010). On the need for embodied and dis-embodied cognition. *Frontiers in Psychology*, **1**(242).
Egorova, N., Shtyrov, Y., & Pulvermüller, F. (2013). Early and parallel processing of pragmatic and semantic information in speech acts: neurophysiological evidence. *Frontiers in Human Neuroscience*, **7**(86).
Evans, V. (2009). *How words mean*. Oxford: Oxford University Press.
Gallese, V. (2007). Before and below 'theory of mind': embodied simulation and the neural correlates of social cognition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **362**(1480), 659–669.

Gallese, V., & Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences*, **2**(12), 493–501.
Garrod, S., & Anderson, A. (1987). Saying what you mean in dialogue: a study in conceptual and semantic co-ordination. *Cognition*, **27**(2), 181–218.
Garrod, S., & Pickering, M. (2004). Why is conversation so easy? *Trends in Cognitive Sciences*, **8**(1), 8–11.
Geeraerts, D. (2010). Recontextualizing grammar: underlying trends in thirty years of Cognitive Linguistics. In E. Tabakowska, M. Choinski, & L. Wiraszka (Eds.), *Cognitive linguistics in action: from theory to application and back* (pp. 71–102). Berlin: Mouton de Gruyter.
Gibbs, R. (2006). Metaphor interpretation as embodied simulation. *Mind & Language*, **21**(3), 434–458.
Glenberg, A., & Gallese, V. (2012). Action-based language: a theory of language acquisition, comprehension, and production. *Cortex*, **48**(7), 905–922.
Glenberg, A., & Kaschak, M. (2002). Grounding language in action. *Psychonomic Bulletin & Review*, **9**(3), 558–565.
Glenberg, A., & Robertson, D. (1999). Indexical understanding of instructions. *Discourse Processes*, **28**(1), 1–26.
Goldman, A. (1992). In defense of the simulation theory. *Mind & language*, **7**(1/2), 104–119.
Gopnik, A. (1995). How to understand beliefs. *Behavioral and Brain Sciences*, **18**(02), 398–400.
Graesser, A., Millis, K., & Zwaan, R. (1997). Discourse comprehension. *Annual Review of Psychology*, **48**(1), 163–189.
Graesser, A., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, **101**(3), 371–395.
Hagoort, P., Hald, L., Bastiaansen, M., & Petersson, K. (2004). Integration of word meaning and world knowledge in language comprehension. *Science*, **304**, 438–441.
Hauk, O., Johnsrude, I., & Pulvermüller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron*, **41**(2), 301–307.
Hauk, O., & Pulvermüller, F. (2004). Neurophysiological distinction of action words in the fronto–central cortex. *Human Brain Mapping*, **21**(3), 191–201.
Havas, D., Glenberg, A., & Rinck, M. (2007). Emotion simulation during language comprehension. *Psychonomic Bulletin & Review*, **14**(3), 436–441.
Holler, J., & Beattie, G. (2003). Pragmatic aspects of representational gestures: Do speakers use them to clarify verbal ambiguity for the listener? *Gesture*, **3**(2), 127–154.
Hostetter, A. (2011). When do gestures communicate? A meta-analysis. *Psychological Bulletin*, **137**(2), 297–315.
Hostetter, A., & Alibali, M. (2008). Visible embodiment: gestures as simulated action. *Psychonomic Bulletin & Review*, **15**(3), 495–514.
Kaschak, M., & Glenberg, A. (2000). Constructing meaning: the role of affordances and grammatical constructions in sentence comprehension. *Journal of Memory and Language*, **43**(3), 508–529.
Kaschak, M., Madden, C., Therriault, D., Yaxley, R., Aveyard, M., Blanchard, A., & Zwaan, R. (2005). Perception of motion affects language processing. *Cognition*, **94**(3), B79–B89.
Kelly, S., Özyürek, A., & Maris, E. (2010). Two sides of the same coin: speech and gesture mutually interact to enhance comprehension. *Psychological Science*, **21**(2), 260–267.
Kendon, A. (1994). Do gestures communicate? A review. *Research on Language and Social Interaction*, **27**(3), 175–200.
Kendon, A. (2004). *Gesture: visible action as utterance*. Cambridge: Cambridge University Press.
Klatzky, R., Pellegrino, J., McCloskey, B., & Doherty, S. (1989). Can you squeeze a tomato? The role of motor representations in semantic sensibility judgments. *Journal of Memory and Language*, **28**(1), 56–77.
Krauss, R., & Fussell, S. (1988). Other-relatedness in language processing: discussion and comments. *Journal of Language and Social Psychology*, **7**(3/4), 263–279.
Krauss, R., & Fussell, S. (1996). Social psychological models of interpersonal communication. In E. T. Higgins & A. Kruglanski (Eds), *Social psychology: a handbook of basic principles* (pp. 655–701). New York: Guilford Press.

Langacker, R. (1987). *Foundations of Cognitive Grammar: theoretical prerequisites. Volume 1*. Stanford: Stanford University Press.
Langacker, R. (1997). The contextual basis of cognitive semantics. In J. Nuyts & E. Pedersen (Eds.), *Language and conceptualization* (pp. 229–252). Cambridge: Cambridge University Press.
Langacker, R. (2008). *Cognitive Grammar: a basic introduction*. Oxford: Oxford University Press.
Linell, P. (1982). *The written language bias in linguistics*. Linkoping: Deptartment of Communication Studies, University of Linkoping. Republished in revised form in 2005 by Routledge (London).
Linell, P. (1998). *Approaching dialogue: talk, interaction and contexts in dialogical perspectives*. Amsterdam: John Benjamins
Linell, P. (2007). Dialogicality in languages, minds and brains: Is there a convergence between dialogism and neuro-biology? *Language Sciences*, **29**(5), 605–620.
Mahon, B., & Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *Journal of Physiology*, **102**(1/3), 59–70.
Marghetis, T., & Bergen, B. (in press). Embodied meaning, inside and out: the coupling of gesture and mental simulation. In Müller, C., Cienki, A., Fricke, E., Ladewig, S., McNeill, D., & Bressem, J. (Eds.), *Body – language – communication: an international handbook on multimodality in human interaction*, Vol. 2. Berlin: Mouton De Gruyter.
Masson, M., Bub, D., & Warren, C. (2008). Kicking calculators: contribution of embodied representations to sentence comprehension. *Journal of Memory and Language*, **59**(3), 256–265.
Matlock, T. (2004). Fictive motion as cognitive simulation. *Memory & Cognition*, **32**(8), 1389–1400.
McNeill, D. (1992). *Hand and mind: what gestures reveal about thought*. Chicago: University of Chicago Press.
Menenti, L., Pickering, M., & Garrod, S. (2012). Toward a neural basis of interactive alignment in conversation. *Frontiers in Human Neuroscience*, **6**(185), 1–9.
Molnar-Szakacs, I., Wu, A., Robles, F., & Iacoboni, M. (2007). Do you see what I mean? Corticospinal excitability during observation of culture-specific gestures. *PLoS One*, **2**(7), e626.
Montgomery, K., Isenberg, N., & Haxby, J. (2007). Communicative hand gestures and object-directed hand movements activated the mirror neuron system. *Social Cognitive and Affective Neuroscience*, **2**(2), 114–122.
Ochs, E., Schegloff, E., & Thompson, S. (1996). *Interaction and grammar*. Cambridge: Cambridge University Press.
Özyürek, A., Willems, R., Kita, S., & Hagoort, P. (2007). On-line integration of semantic information from speech and gesture: insights from event-related brain potentials. *Journal of Cognitive Neuroscience*, **19**(4), 605–616.
Parker, I. (1992). *Discourse dynamics: critical analysis for social and individual psychology*. London: Routledge.
Pickering, M., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, **27**(02), 169–190.
Pickering, M., & Garrod, S. (2009). Prediction and embodiment in dialogue. *European Journal of Social Psychology*, **39**(7), 1162–1168.
Pulvermüller, F., Mohr, B., & Schleichert, H. (1999). Semantic or lexico-syntactic factors: What determines word-class specific activity in the human brain? *Neuroscience Letters*, **275**(2), 81–84.
Rączaszek-Leonardi, J. (2009). Symbols as constraints: the structuring role of dynamics and self-organization in natural language. *Pragmatics & Cognition*, **17**(3), 657–676.
Richardson, D., Spivey, M., Barsalou, L., & McRae, K. (2003). Spatial representations activated during real-time comprehension of verbs. *Cognitive Science*, **27**(5), 767–780.
Robinson, E. (2000). The cognitive foundations of pragmatic principles: implications for theories of linguistic and cognitive representation. In J. Nuyts & E. Pedersen (Eds.), *Language and conceptualization* (pp. 253–271). Cambridge: Cambridge University Press.

Sato, M., Mengarelli, M., Riggio, L., Gallese, V., & Buccino, G. (2008). Task related modulation of the motor system during language processing. *Brain and Language*, **105**(2), 83–90.

Shannon, C., & Weaver, W. (1948). *The mathematical theory of communication*. Urbana: University of Illinois Press.

Skipper, J., Goldin-Meadow, S., Nusbaum, H., & Small, S. (2007). Speech-associated gestures, Broca's area, and the human mirror system. *Brain and Language*, **101**(3), 260–277.

Speer, N., Zacks, J., & Reynolds, J. (2007). Human brain activity time-locked to narrative event boundaries. *Psychological Science*, **18**(5), 449–455.

Sperber, D., & Wilson, D. (1986). *Relevance: communication and cognition*. Cambridge, MA: Harvard University Press.

Stanfield, R., & Zwaan, R. (2001). The effect of implied orientation derived from verbal context on picture recognition. *Psychological Science*, **12**(2), 153–156.

Stolorow, R., & Atwood, G. (1992). *Contexts of being: the intersubjective foundations of psychological life*. Hillsdale: Analytic Press.

Swerts, M., & Geluykens, R. (1994). Prosody as a marker of information flow in spoken discourse. *Language and Speech*, **37**(1), 21–43.

Tettamanti, M., Buccino, G., Saccuman, M., Gallese, V., Danna, M., Scifo, P., Fazio, F., Rizzolatti, G., Cappa, S. F., & Perani, D. (2005). Listening to action-related sentences activates fronto-parietal motor circuits. *Journal of Cognitive Neuroscience*, **17**(2), 273–281.

Thibault, P. (2005). The interpersonal gateway to the meaning of mind: unifying the inter- and intraorganism perspective on language. In R. Hasan, , C. Matthiessen, & J. Webster (Eds.), *Continuing discourse on language: a functional perspective* (pp. 117–156). London: Equinox.

Van Berkum, J., Van Den Brink, D., Tesink, C., Kos, M., & Hagoort, P. (2008). The neural integration of speaker and message. *Journal of Cognitive Neuroscience*, **20**(4), 580–591.

Van Dam, W., Rueschemeyer, S., Lindemann, O., & Bekkering, H. (2010). Context effects in embodied lexical-semantic processing. *Frontiers in Psychology*, **1**(150).

Van Dijk, T., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.

Van Elk, M., Slors, M., & Bekkering, H. (2010). Embodied language comprehension requires an enactivist paradigm of cognition. *Frontiers in Psychology*, **1**(234).

Varela, F., Thompson, E., & Rosch, E. (1991). *The embodied mind: cognitive science and human experience*. Cambridge, MA: MIT press.

Vigliocco, G., Warren, J., Siri, S., Arciuli, J., Scott, S., & Wise, R. (2006). The role of semantics and grammatical class in the neural representation of words. *Cerebral Cortex*, **16**(12), 1790–1796.

Wallentin, M., Nielsen, A. H., Vuust, P., Dohn, A., Roepstorff, A., & Lund, T. E. (2011). BOLD response to motion verbs in left posterior middle temporal gyrus during story comprehension. *Brain and Language*, **119**(3), 221–225.

Willems, R., Benn, Y., Hagoort, P., Toni, I., & Varley, R. (2011). Communicating without a functioning language system: implications for the role of language in mentalizing. *Neuropsychologia*, **49**(11), 3130–3135.

Willems, R., de Boer, M., de Ruiter, J., Noordzij, M., Hagoort, P., & Toni, I. (2010). A dissociation between linguistic and communicative abilities in the human brain. *Psychological Science*, **21**(1), 8–14.

Willems, R., & Francken, J. (2012). Embodied cognition: taking the next step. *Frontiers in Psychology*, **3**(582).

Willems, R., & Hagoort, P. (2007). Neural evidence for the interplay between language, gesture, and action: a review. *Brain and Language*, **101**(3), 278–289.

Willems, R., & Varley, R. (2010). Neural insights into the relation between language and communication. *Frontiers in Human Neuroscience*, **4**(203).

Wilson, A., & Golonka, S. (2013). Embodied cognition is not what you think it is. *Frontiers in Psychology*, **4**(58).

Winter, B., & Bergen, B. (2012). Language comprehenders represent object distance both visually and auditorily. *Language and Cognition*, **4**, 1–16.

Yaxley, R., & Zwaan, R. (2007). Simulating visibility during language comprehension. *Cognition*, **105**(1), 229–236.
Zwaan, R. (2003). The immersed experiencer: toward an embodied theory of language comprehension. *Psychology of Learning and Motivation*, **44**, 35–62.
Zwaan, R. (2009). Mental simulation in language comprehension and social cognition. *European Journal of Social Psychology*, **39**(7), 1142–1150.
Zwaan, R. (2014). Embodiment and language comprehension: reframing the discussion. *Trends in Cognitive Sciences* **18**(5), 229–234.
Zwaan, R., Madden, C., Yaxley, R., & Aveyard, M. (2004). Moving words: dynamic representations in language comprehension. *Cognitive Science*, **28**(4), 611–619.
Zwaan, R., Magliano, J., & Graesser, A. (1995). Dimensions of situation model construction in narrative comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **21**(2), 386–397.
Zwaan, R., & Radvansky, G. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, **123**(2), 162–185.
Zwaan, R., Stanfield, R., & Yaxley, R. (2002). Language comprehenders mentally represent the shapes of objects. *Psychological Science*, **13**(2), 168–171.
Zwaan, R., & Taylor, L. (2006). Seeing, acting, understanding: motor resonance in language comprehension. *Journal of Experimental Psychology: General*, **135**(1), 1–11.

https://doi.org/10.1017/langcog.2014.25 Published online by Cambridge University Press