



THE MANY WHYS OF THE PSYCHOLINGUISTICS OF GESTURE

Jakub Jehlička¹

Breckinridge Church, R., Alibali, M. W., & Kelly, S. (Eds.) (2017). *Why Gesture? How the Hands Function in Speaking, Thinking and Communicating*. Amsterdam: John Benjamins.

“Why gesture?” — this question is in the title of the seventh volume of John Benjamins book series dedicated to gesture studies. With two monographs (*Gesturecraft* by Jurgen Streeck (2009) and *Elements of Meaning in Gesture* by Genevieve Calbris (2011) and four edited volumes (Cienki & Müller, 2008; McNeill, Duncan, Cassell, & Levy, 2007; Pika & Liebal, 2012; Stam & Ishino, 2011) published so far, the “Gesture studies” book series (edited by Adam Kendon) now constitutes a representative body of texts reflecting the multifaceted nature of this dynamic area of linguistics, that instead of traditional view characterized by the label “non-verbal communication” approaches language in an integrative way, viewing it as multimodal and semiotically diverse communication system.

The latest contribution to the series is a collection of 18 chapters edited by Ruth Breckinridge Church (Northeastern Illinois University), Martha Alibali (University of Wisconsin-Madison) and Spencer Kelly (Colgate University) that focus on hand gestures’ role in “speaking, thinking and communicating” as the book’s subtitle tells us. The title question is a bit enigmatic (certainly intentionally), due to the polysemy of the English noun/verb *gesture*: one may wonder if it is supposed to mean *why do we gesture?* or perhaps *why, of all possible means of nonverbal communication, humans favour gesture?*

In the introductory chapter, the editors clarify, that what they are actually asking is “What are the many and varied effects of gesture for producers and observers, and how can one account for these effects in neural, cognitive and social terms?” (p. 5). It is not only the *effects*, but also the *causes* of gesturing what is in focus — the editors set the scene of this book by reflecting Aristotle’s framework for describing complex behaviours (*efficient cause* and *final cause* as described in his *Physics* and *Metaphysics*) The editors further explain that the book’s focus is on two areas: *Mechanisms* of cognitive processing of gestures (psycholinguistics of gesture, describing the *efficient cause* of gesture) and *functions* of gestures in human communication (i.e., studies analysing the *final cause* of gesture).

Another way of looking at this book is through the lens of its contributors: we can see that the majority of the authors are currently based in the United States or Canada (11 out of 18). To be more specific, most of the contributors are affiliated to universities in the Midwest of the United States, an area with a symbolic epicentre at the University of Chicago, which, thanks to David McNeill and Susan Goldin-Meadow,

1 This work was supported by the European Regional Development Fund-Project “Creativity and Adaptability as Conditions of the Success of Europe in an Interrelated World” (No. CZ.02.1.01/0.0/0.0/16_019/0000734).

can be considered one of the points of origin of the discipline, and, today, a co-speech gesture research hub in North America. *Why Gesture?* thus also represents the current North American perspective on the study of co-speech gestures.

In this review, I will take a closer look only on a few of the chapters — the ones that may be subsumed under the label of the *psycholinguistics of co-speech gesture*. More specifically, I will focus on the contributions discussing the models of co-speech gesture cognitive processing and also sharing another common denominator — reflection of David McNeill’s idea of representational duality of gesture and speech (1992, 2005, *inter alia*).

The 18 chapters are divided into three sections: “The function of gesture production for language”, “The function of gesture comprehension” and “Why gesture? Some theoretical implications”. The first section starts with a chapter by Martha W. Alibali, Amelia Yeo, Autumn B. Hostetter and Sotaro Kita titled “Representational gestures help speakers package information for speaking”. Here, the authors revisit Kita’s (2000) *Information Packaging Hypothesis* (IPH), one of the early psycholinguistic models of gesture production. As was a considerable part of the research in this area, Kita too was inspired by McNeill’s ideas of what could be called the *cognition of gesture* — that in turn had its roots in Vygotsky’s views on mental representation (Vygotsky, 1986), for which an idea of *dual* nature of thought was crucial. McNeill claims that when producing co-speech gestures, speakers dynamically combine two opposing semiotic modes of conceptual representation, i.e. *imagistic* and *linguistic* (or propositional), into so-called *Growth Points*, that have been characterized as the basic organizational units of thought or of “linguistic embodiment” (see McNeill’s chapter in the reviewed volume — which will not be discussed here).

In IPH, Kita similarly distinguishes between *analytical thinking* and *spatio-motoric thinking*. Both modes of thought play a role in the packaging of information for speaking — while analytical thinking is the primary mode behind the information organization in speech, production of gestures may influence the way we speak by activating the spatio-motoric thinking to a greater extent. Such a view echoes the neo-Whorfian inspiration in psycholinguistic research which was very much in vogue in the late 1990s and early 2000s (Gumperz & Levinson, 1996). In this chapter, the authors review the evidence for IPH that has accumulated since its proposal. In general, the evidence supports the assumption of bidirectional influence between gesture and speech, which was put forward by Kita in his collaborative works with Aslı Özyürek. Their *Interface Hypothesis* (IH, Kita & Özyürek, 2003) was designed as an add-on to the famous Levelt’s “blueprint” model of speech production (Levelt, 1989). IH and IPH are compatible models — IH describes the inner workings of conceptualization, which is a part of the complex model of gesture-speech integration captured by IPH.

The following chapter, “Function and processing of gesture in the context of language” by Aslı Özyürek, is closely linked to the previous one. Özyürek (currently leading the Multimodal Language and Cognition lab at the Max Planck Institute for Psycholinguistics in Nijmegen) presents another body of evidence along the same lines. She reviews the crosslinguistic evidence — the main area of the IPH-related





research has been crosslinguistic comparison — of how gesture affects speech². This evidence is based on the analyses of English, Japanese and Turkish gesture production, that were mostly compared in terms of the use of gesture accompanying motion event expressions. Following Leonard Talmy's semantic typology, built upon different patterns of framing the motion events (so-called *satellite-framing* and *verb-framing languages*; see Talmy, 1985), these studies (e.g. Kita et al., 2007; Özyürek, Kita, Allen, Furman, & Brown, 2005) have shown that the preferences for highlighting different aspects of motion event internal structure (i.e. focus on *path* or *manner* of the motion or both at the same time) in speech also lead to different use of iconic gestures (again, highlighting the above mentioned aspects of event structure). Along the lines of the neo-Whorfian approach, Özyürek points out that when “we consider many languages [...] we see that gesture reflects differences in the way thoughts are linguistically organized in speakers” (p. 54).

J. P. de Ruiter, author of a chapter titled “The asymmetric redundancy of gesture and speech”, was, like Kita and Özyürek, formerly associated with the Nijmegen MPI for Psycholinguistics. Now leading the Human Interaction Lab at Tufts University, he combines seemingly incompatible methodologies — conversation analysis and experimental methods in his current research. Yet, in the present chapter, he too turns back to an older theory of his — the *Sketch Model* (de Ruiter, 2000) — revisiting it according to more recent evidence. Inspired by McNeill's idea of dual nature of conceptual representation on the one hand and by Levelt's model of speech production on the other, the Sketch Model, represents what de Ruiter calls the *Postcard Architecture* (de Ruiter, 2007) of a model hierarchy in which the “information to be communicated is dispatched into gesture and speech channels by a central process”³, but, like a postcard, having two sides: “imagistic” and “propositional”. In terms of Levelt's model, the separation of the two modes of representation occurs in the *Conceptualizer/Communication Planner* and they are further processed by separate nodes of *Action Generation* (gesture) and *Formulator* (speech).

In the present chapter, de Ruiter acknowledges that the idea of imagistic and propositional information discrete complementarity does not seem to be supported by the empirical findings (including the same kind of evidence reviewed by Özyürek) supporting the idea that gesture is to a significant extent and at various levels dependent on speech, and hence he proposes a modification of the original model. As the aggregated evidence rules out the assumptions that gestural and linguistic part of

2 In IPH/IH the link between gesture and speech is bidirectional, allowed by separation of the *Action Generator* and *Message Generator* nodes in the model architecture.

3 De Ruiter (2007) identifies three general ways in which processes of speech and gesture production are organized in psycholinguistic models, apart from *Postcard Architecture*, he distinguishes *Window Architecture* — assuming that (iconic) gestures are a direct manifestation of thought, unconstrained by linguistic conceptualization, thus providing “a window to the mind” — and *Language Architecture* — that, like the *Postcard Architecture*, assumes that speech and gesture are both part of communication intention, but it “differ[s] in the way the semantic synchronization between language and gesture is achieved” (de Ruiter, 2007, p. 36).

the message are two separate channels, de Ruiter instead proposes *Asymmetric Redundancy (AR) Sketch Model*. The input of *Action Generation* and *Formulator* is no longer complementary, but redundant — the imagistic and propositional representations are interlinked. The relation between the two modes of representations is asymmetric: in the AR Sketch Model, propositional content is fed directly to the *Formulator*, whereas imagistic content is constrained by the activation of specific propositional content before proceeding the *Action Generator*.

The last chapter I will stop by here is by Spencer D. Kelly (“Exploring the boundaries of gesture-speech integration during language comprehension”), who presents yet another model of gesture-speech processing, this time from a comprehension/perception perspective. Like IH, IPH and AR-Sketch Model, this model, referred to as *Integrated Systems Hypothesis (ISH)*, is seen as a development of the McNeillian idea that “gestures imagistically index what is novel or relevant in a spoken utterance within a given context” (p. 244) — elaborating on it by asking “what aspects of spoken language are open to this “news” delivered through gestures?” (ibid.).

Developed more recently (Kelly, Özyürek & Maris, 2010; Kelly, 2017), ISH does not exhibit the hierarchical structure of classic psycholinguistic models, unlike the previous accounts. ISH aims at capturing the interplay of gesture and speech at different structural levels: semantic, phonological, syntactic and pragmatic. Kelly provides a review of empirical evidence indicating, that apart from relatively well-researched semantic level, integrated processing of gesture and speech during language comprehension occurs most strongly at the suprasegmental phonetic level (i.e. prosody) and in understanding pragmatics (e. g. when gesture discriminates between competing speech acts). At the segmental phonetic and syntactic levels, evidence is not yet conclusive, or it suggests a lesser role of gesture. As for semantics, the evidence points to deeper integration in case of more concrete (i.e. more directly embodied) concepts.

ISH diverges from the previous models not only by actually transcending the domain of semantics (focus on semantics often entails, in effect, resorting only to iconic gestures), but also in asking how exactly the integration functions at various temporal levels (during on-line processing and from long-term and developmental perspectives).

The first three chapters reviewed here present revisions of psycholinguistic models originating at least 20 years ago. Since then, research of cognitive processing of co-speech gestures took a giant leap forward: production studies have benefited from a growing number of multimodal corpus data with a growing emphasis on the degree of ecological validity, comprehension research has witnessed development of new and more reliable experimental methods. Generally, implementation of neurocognitive methods and tools and motion capture technologies has provided a considerable boost. Discussions of the older models presented here reflect this development — after all, the authors have played a pivotal role in it. However, in some respects, the general framing of these chapters is still deeply embedded in the theoretical frameworks of the early 2000s or even earlier.

Although the later modifications of Levelt’s model capture production as well comprehension, IH, IPH and AR-Sketch model are only centred on the production mechanisms. From today’s perspective, this seems to be inadequate: language pro-





duction does not occur in a vacuum, and a truly robust model should be built in the context of a more complex communication situation with both speaker and listener engaged in interaction. In such context, it is not a model based on a single speaker's production/comprehension of a single word, but rather a model based on a *turn-sequence* that would be more plausible.

The models discussed here also are (more or less) explicitly limited to iconic gestures only. But is drawing such a clear division line between iconic and non-iconic or “representational” and other gestures justified? After all, how typical are clear-cut iconic gestures for natural communication? If iconicity is rather one of the dimensions of semiotically diverse and principally polyfunctional gestures, why should we construct a separate model for iconic gestures processing while ignoring the other dimensions?

Kelly's holistic and integrative theory, on the other hand, takes into account all dimensions of co-speech gestures by widening the scope beyond semantic processing. With its focus on different structural levels, ISH is also no longer fixed on a single word as a basic structural unit for analysing the gesture-speech integration. Reflecting the neurological evidence, Kelly even points at the limitations and usefulness of generalizing of the gesture-speech integration process in terms of “conceptual blending of two different representational formats” (p. 249) and directs the attention of gesture researchers to probabilistic modelling of neural activity as a new way of understanding the processes of multimodal communication.

All in all, *Why Gesture?* is a valuable volume. Although some of the chapters may be outdated in their premises, they nevertheless provide a worthy insight as to the paths research in this field has taken over the last two decades. Together with the rest of the book, they offer a good and representative overview of the current state of the thriving and dynamic area of gesture studies, perhaps still not providing definitive answers to all the *whys* but certainly inspiring new ideas about *how* we use, process, and understand co-speech gestures.

REFERENCES

- Calbris, G. (2011). *Elements of Meaning in Gesture*. Amsterdam: John Benjamins.
- Cienki, A. J., & Müller, C. (Eds.). (2008). *Metaphor and Gesture*. Amsterdam: John Benjamins.
- de Ruiter, J. P. (2000). The production of gesture and speech. In D. McNeill, *Language and Gesture* (pp. 248–311). Cambridge: Cambridge University Press.
- deRuiter, J. P. (2007). Postcards from the mind: The relationship between speech, imagistic gesture, and thought. *Gesture*, 7(1), 21–38.
- Gumperz, J. J., & Levinson, S. C. (Eds.). (1996). *Rethinking Linguistic Relativity*. Cambridge: University Press.
- Kelly, S. D., 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.
- Kita, S. (2000). How representational gestures help speaking. In D. McNeill (Ed.), *Language and Gesture* (pp. 162–185). Cambridge: Cambridge University Press.
- Kita, S., & Özyürek, A. (2003). What does cross-linguistic variation in semantic coordination of speech and gesture reveal?: Evidence for an interface representation of spatial thinking and speaking. *Journal of Memory and Language*, 48(1), 16–32.

- Kita, S., Özyürek, A., Allen, S., Brown, A., Furman, R., & Ishizuka, T. (2007). Relations between syntactic encoding and co-speech gestures: Implications for a model of speech and gesture production. *Language and Cognitive Processes*, 22(8), 1212–1236.
- Levelt, W. J. M. (1989). *Speaking: From Intention to Articulation*. Cambridge, MA: The MIT Press.
- McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago, IL: University of Chicago Press.
- McNeill, D. (2005). *Gesture and Thought*. Chicago, IL: University of Chicago Press.
- McNeill, D., Duncan, S. D., Cassell, J., & Levy, E. T. (Eds.). (2007). *Gesture and the dynamic dimension of language: Essays in honor of David McNeill*. Amsterdam: J. Benjamins.
- Özyürek, A., Kita, S., Allen, S., Furman, R., & Brown, A. (2005). How does linguistic framing of events influence co-speech gestures?: Insights from crosslinguistic variations and similarities. *Gesture*, 5(1–2), 219–240.
- Pika, S., & Liebal, K. (Eds.). (2012). *Developments in Primate Gesture Research*. Amsterdam: John.
- Stam, G., & Ishino, M. (Eds.). (2011). *Integrating Gestures: The Interdisciplinary Nature of Gesture*. Amsterdam: John Benjamins.
- Streeck, J. (2009). *Gesturecraft: The Manu-facture of Meaning*. Amsterdam: John Benjamins.
- Talmy, L. (1985). Lexicalization patterns: Semantic structure in lexical forms. In T. Shopen (Ed.), *Language typology and syntactic description, Vol.3, Grammatical categories and the lexicon* (pp. 57–149). Cambridge: Cambridge University Press.
- Vygotsky, L. S. (1986). *Thought and Language*. Cambridge, MA: MIT Press.



Jakub Jehlička | Department of Linguistics, Faculty of Arts, Charles University / Department of Linguistics and Cognitive Semiotics, RWTH Aachen
 <jakub.jehlicka@ff.cuni.cz>