

Metaphor and Gesture Alan Cienki and Cornelia Muller (eds) (2008)

Reviewed by Radan Martinec

The book to be reviewed is a collection of contributions from scholars of gestures that co-occur with speech, and focuses on gestures that function metaphorically. The chapters are based on papers that were presented at a cognitive linguistics conference in Spain in 2003 and the contributors are from Europe and North America. The chapters are followed by brief comments from scholars whose work has touched on the issues raised in the volume. Lakoff and Johnson's (e.g. 1980) cognitive theory of metaphor, in which metaphor is conceptualized as a mapping between a source domain (usually more concrete) and a target domain (usually more abstract), is used throughout the contributions. McNeill's (e.g. 1992, 2005) theory, according to which co-verbal gestures are non-systematic (without regular correspondences between meanings and forms) context-sensitive (specific to particular contexts), and idiosyncratic (specific to an individual), inspires the approach taken to gestures in all but one chapter. I will first summarize and evaluate the chapter contributions and follow with a comment on McNeill's gesture theory. For reasons of space, the comments in the second part of the reviewed book will be referred to only every so often, as they impact on my discussion and evaluation of the main part of the book.

Cienki

Cienki's chapter is a survey of issues regarding metaphor in language and gestures. He points out that metaphors lie on a scale of conventionality and creativity, both on the conceptual and expression side. For example, 'life is

Affiliation

IKONA Research and Consulting, Arizona, USA.
email: radan.martinec@gmail.com

a banana' is a more creative and less conventionalized conceptual metaphor than 'life is a journey'. On the other hand, 'my life is going nowhere' is a more conventional way of expressing the metaphor 'life is a journey' than 'he skateboarded his way through life'. The same is said to apply to co-verbal gestures, although it is not made clear how the metaphorical expressions can be more or less conventionalized if the gestures are said to be 'spontaneous' (p. 10), and therefore presumably always creative and not obeying conventions.

There is a useful discussion of the possible relationships between metaphors in speech and co-verbal gestures. The same metaphoric mapping can, for example, be realized in both speech and gesture concurrently, as when a gesture gives a more detailed representation of a more general meaning expressed in a lexical item. But gestures and words can also serve different expressive functions at the same time – for example, when the gesture expresses the source and the speech the target domain of a metaphorical mapping. Gestures and words can also realize different sources for the same target domains, and so on. This is a welcome discussion of the gesture-speech relations in the context of metaphor, and one can only wish that it were formalized in a taxonomy or a system network of the kind used in systemic-functional linguistics with rules specifying the realizations of the various options. Much of what Cienki says about gesture and speech relations appears to be amenable to modeling by the same or slightly altered network for image-text relations developed in Martinec and Salway (2005), where an application of the system to speech-gesture relations was suggested.

Cienki makes some proposals for how the study of gesture could benefit the study of metaphor, of which the most interesting I find the idea elaborated in more detail in some of the following chapters, that evidence from the study of gesture can avoid the circularity that the linguistic theory of metaphor has been charged with. This is followed by several questions that the study of gesture raises for conceptual metaphor theory. In the context of this review, the most interesting issue that Cienki raises is to do with the consciousness of the speakers of metaphorical gestures – should only such gestures be analyzed that speakers are conscious of, or should any gestures be considered that may be potentially metaphoric in nature? This hinges on the issue of how conscious speakers are of co-verbal gestures in general. In the closing section of this review, I will argue that speakers are generally unconscious of co-verbal gestures, and thus, if one analyzed only such metaphors that speakers are conscious of, one would not get very far.

Calbris

Calbris takes a semiotic approach to gestures, building on her previous work where she identified abstract concepts and their realizations, such as evolution

in time being realized by a succession of looping gestures (e.g. Calbris 1990). In her chapter, she focuses on the metaphorical realizations of four semantic axes onto one transversal axis. The symbolic axes are a value axis, a spatio-temporal axis, a logico-temporal axis and an axis depicting an evolving process. The transversal axis is a physical axis running from the left to the right of the body, and Calbris is concerned with how gestures translate abstract, symbolic meanings into concrete, perceivable representations by means of locating them on the physical axis. The mappings are metaphorical because the locations on the more concrete axis stand for the more abstract concepts.

For example, when representing values, ideas generally perceived as negative, like 'regression', 'reduction' and 'deficit', are located on the left, whereas positive ideas like 'increase' are on the right. So Lionel Jospin, the Prime Minister at the time, whose televised interviews constitute the data, says 'we continue to increase public expenditure' while moving his right hand to the right, and 'we continue to reduce our deficits' while moving his left hand to the left (p. 43). As for the spatio-temporal axis, a path in space or time is depicted by a left-right movement, most likely reflecting the influence of writing in western cultures. For the same reason, cause on the logico-temporal axis, which usually precedes effect, is located on the left and the effect on the right of the transversal axis.

Calbris' analysis of both gesture meanings and forms is systematic, and is the only one in this collection that adopts a semiotic approach, with the rest following a more eclectic, mostly cognitivist route, inspired by McNeill (e.g. 1992, 2005). One thing that is lacking in Calbris' otherwise inspiring chapter is a more diverse data set. Only one speaker's gestures are analyzed and it is perhaps somewhat difficult to make generalizations about metaphorical gestures produced by one speaker, in one type of context (televised political interviews), across other speakers and contexts. Could she perhaps be dealing with the speaker's gestural idiolect? This appears especially important in light of all the other contributors following, more or less closely, McNeill's argument that co-verbal gestures are idiosyncratic and context-specific.

Williams

Williams analyzes the use of gestures in an instructional setting (a teacher explaining how to read the clock to primary school students) to argue that gestures help the speaker encode and addressee to understand the conceptual mappings between the source and target domains in conceptual metaphors. On the basis of such realizations in gesture, the bridge between the two domains can more easily be inferred by the addressee. This would be difficult to do, or in any case less efficient, if only language was used. What follows is that conceptual metaphor has multimodal realizations.

Strictly speaking, Williams draws on a theory of conceptual blending (see e.g. Fauconnier and Turner, 2002), which is slightly different from metaphor in that while conceptual metaphor analyses typically focus on entrenched patterns of cross-domain mappings, conceptual blending has to do with producing novel meanings. Both theories are however based in the same concept of a mapping between a source domain and a target domain.

Williams' use of gestures to make explicit the cross-domain mappings between quadrants of a circle introduced earlier in a class, and those of the clock face introduced later is quite convincing and commendably explicit. It thus makes a valuable contribution to the use of gestures related to this aspect of the conceptual metaphor theory. His distinction between metaphor and blending, however, could have been made more clearly, especially because some of the later chapters argue for the role of gestures in realizing instantial, or 'unexpected' (McNeill), metaphors. It is thus an open question if the introduction of conceptual blending in a volume on metaphor is truly necessary. At the very least, the advantages of using conceptual blending over metaphor would have deserved a more thorough discussion.

Nunez

While Williams' chapter was about the role of gestures in the teaching of a conceptual system, Nunez's chapter deals with the role of gestural representation in providing evidence for the organization itself of the conceptual system of mathematics. Following on the work he produced in collaboration with Lakoff (e.g. Lakoff and Nunez, 2000), he argues that, because of the obvious circularity, it does not make much sense to explain mathematics by reasoning derived from itself. Instead, the most fundamental concepts of often very abstract sciences like math come from everyday, embodied concepts and abstractions derived from these by the process of metaphor.

The evidence he provides for the organization of the conceptual system of mathematics consists in analyses of how a professional mathematician's gestures dynamically realize metaphorical concepts that concurrently occur in his mathematical explanations. Nunez's chapter is the first of several in this volume that follow the argument that co-verbal gestures provide evidence of the psychological reality of the metaphors that are concurrently realized in speech. This argument seems to have two aspects to it, one of which regards individual gestures and the other recurring gestures of the same or similar form over a stretch of discourse. As for the individual gestures, the purportedly direct realizations of their metaphorical meanings in forms makes it possible for them to represent the spatio-metaphorical thinking that is at the core of many metaphors. The presence of similar gestures over a stretch of

discourse realizes a ‘catchment’, which is a stretch of discourse characterizing a single textual instance. This instantial, dynamic quality of catchments is said to also argue for the metaphorical concepts realized in it to be live (see below for a further discussion of the two characteristics of gestures).

According to Nunez, who unfortunately does not provide data transcripts, the presence of co-verbal gestures over a stretch of discourse provides evidence of the psychological reality of the metaphors that are concurrently realized in speech. The reasoning is that it is difficult to decide on the basis of linguistic evidence alone whether metaphors are ‘dead’, and thus not processed as metaphors by the speakers, or whether they are still ‘live’. If one can observe the gestures’ forms iconically and repeatedly realizing the objects or actions that the abstract, metaphorical concepts are based on, this provides evidence that the metaphors are still active in the speakers’ minds. I do find this argument fairly convincing, although I do not believe that one needs a purely instantial gesture theory to support it. Phasal analysis (see below) in systemic-functional linguistics (e.g. Gregory 1985a, b, 1988) maps out dynamic, linear sequences of units retrieved from the linguistic system that are particular to single texts.

Mittelberg

Mittelberg’s chapter combines the conceptual theory of metaphor and McNeill’s gesture theory with Peircean semiotics. Metaphoric gestures have so far been considered to be based on realizations of abstract meanings in iconic forms, where iconicity has been conceptualized as a broad similitude of the gestures’ forms to some concrete object (the source domain), which is mapped by the process of metaphoric mapping, analyzed in some detail in Williams’ chapter, onto the target domain. One of the two main contributions of Mittelberg’s chapter is to show how the mappings work for a more detailed classification of icons that she borrows from Peirce (e.g. 1955), viz. image, diagram and metaphor. While analyzing the use of metaphoric gestures in a class performance of a linguistics professor, she finds, fairly unsurprisingly, that, because of their ability, or affordance, to iconically represent objects and actions, gestures tend to be better than language at representing both the pictorial and diagrammatic kinds of the source domains, and thus also the meanings of their metaphorical target domains.

Another contribution, which appears more substantial, is that iconicity and metaphor are not enough to account for the realization of metaphors in gestures, and that an introduction of metonymy, or Peirce’s ‘indexicality based on contiguity’, should also be part of the picture. Her reasoning for the inclusion of metonymy is based on Peirce’s concept of the ‘ground’, which is that part of the sign that is actually relevant to the representation of the object.

It is certainly true that not all aspects of a gesture are relevant to realizing a concept, literal or metaphorical. And it is equally true that only certain aspects of a concept are realized in a gesture, as in Mittelberg's example of a cupped hand realizing the well-known conduit metaphor. In order to understand the metaphor, one has to infer from the cupped palm the object that is imaginarily placed in it, which is some aspect of communication (e.g. a single word or a whole discourse). So, quite rightly, the perceivable gesture form itself only represents one part of the overall metaphorical meaning. The figure of speech or, more broadly, the cognitive concept of metonymy, in which a part stands for the whole, must thus be drawn on by the addressee for the gesture to be correctly interpreted.

McNeill

McNeill's chapter discusses gesture metaphors that he terms 'unexpected'. These are characterized by both gesture forms' and meanings' being created on the spur of the moment and they contrast with 'expected' metaphoric gestures, whose meanings are conventionalized, as evidenced mainly by verbal examples of the same source-target domain mappings, and whose forms also have a degree of conventionalization (which however is dismissed as unimportant). The conduit metaphor (see above) is an example of such a conventionalized metaphoric gesture. An example of an unexpected metaphor is a person reaching up to a certain height with one hand, while keeping the other hand at the waist, a two-hand gesture that depicts respectively the location of a windowsill and of the pavement in a narrative scenario, to realize the meaning that something is not achievable.

While expected metaphorical gestures are supposedly readily recognized by native speakers, unexpected metaphors are not, and one has to take the context of the surrounding discourse into account to realize that such a gesture is indeed metaphoric. This leads to an explanation for the reason of these gestures' existence, which is that they provide a bridge between the growth point and the following discourse. The growth point is analogous to Vygotsky's (1987) psychological predicate, it seems to be related to Halliday's (e.g. 1994) new information, and is said by McNeill to consist of both a language and an image part. It is said to be developed in a stretch of discourse that follows and gestures play the main role in its development. Unexpectedly metaphorical gestures keep recurring over the stretch of discourse that follows the growth point until the growth point's development is exhausted and the topic switches to another growth point. Discourse thus unfolds as a succession of growth points followed by their developments. The recurring gestures are said to form a 'catchment', which is basically a stretch of consis-

tent visuospatial imagery realized by gestures related by some similarity of form (see McNeill 2005: 117).

I find it difficult to understand why unexpected metaphors in particular should form the bridge between the growth point and the catchment, if McNeill says (p. 167) that expected metaphors function the same way. It seems to me that perhaps this has to do with McNeill's (1992, 2005) argument that since the growth point and catchment are particular to single textual instances, the gestures and the meanings that these realize must also be instantial. As the above comment regarding Gregory's phase makes clear, however, this need not necessarily be the case.

Montredon *et al.*

Montredon *et al.*'s chapter takes inspiration from McNeill's concepts of the growth point and catchment and presents an analysis of Jacques Derrida's co-verbal gestures during the conclusion of an interview on a French TV program. It relates the two concepts to Tuite's (1993) theory of verbal and non-verbal production, heavily influenced by McNeill and to some extent also by Kendon (1972, 1980), according to which speech production starts with a combination of speech and image and is driven by synchronized rhythmic pulses in intonation and gesture. Tuite's combination of imagery and language is seen as analogous to McNeill's growth point and Tuite's rhythmic pulses are said to spread over catchments.

Montredon *et al.*'s analysis seems, in some respects, quite similar to the systemic-functional phasal analysis mentioned above, applied to multimodal texts as in for example Martinec (1998, 2000a, 2000b). There, the phase is also seen as correlating with rhythmic waves, and the growth points that Montredon *et al.* identify seem to correspond to the beginnings of sixth level wave phases, which tend to introduce new topics, and are realized by the most prominent rhythmic accents (see also Martinec 2002). The main differences are that, in Martinec's model, rhythm is seen as separate from intonation, with its own units and prominences, and the multimodal aspect of phases regards not only gestures but also other kinds of embodied action. Martinec's rhythm model is also considerably more elaborate and systematic compared with the rather cursory references to rhythm in Tuite (1993), who is drawing on Kendon's (1972, 1980) rhythmic hierarchy that conflates intonation and rhythm.

There is a lack of any attempt to measure the distances between the intonational prominences in Montredon *et al.*, which of course makes it hard for the authors to make a convincing argument for their regularly timed, rhythmical occurrences. Finally, since Montredon *et al.* perform a detailed analysis of the three final catchments in Derrida's interview excerpt, it may be worth men-

tioning that I have found their transcription method difficult to make sense of, because the synchronization points between gestures and speech cannot easily be determined and the authors' description does not make things any clearer. A transcription method of the rhythmic structuring of various semi-otic modes and their synchronization along the lines of Martinec (2000b) or similar would have been much more intelligible.

Parrill

The aim of Parrill's chapter was to empirically test one aspect of McNeill's gesture theory, namely that co-verbal gestures differ from gestures that occur on their own and have a definite meaning (referred to as 'emblems' by Ekman and Friesen (1969) and others), by having no standards of form recognized by native speakers. In particular, she chose to test whether a metaphorical gesture which appears frequently with speech and is performed more or less the same by various speakers has, after all, a conventionalized form or not. The gesture consists of the movement of an arm with the palm held up, extended and horizontal to the ground, metaphorically presenting an item of information by a speaker to a hearer (it is thus rather similar to, although not identical with, the 'conduit metaphor' gesture mentioned above).

According to Parrill, if the gesture were to be proven to be conventionalized, it would mean that it has lost the quality of realizing a live metaphor, which has instead become dead. The reasoning is that if speakers did recognize such a standard, this would mean that the gesture form would be motivated by a social agreement and not by spatio-visual thinking – the assumption being that the realization of spatio-visual thinking, and presumably the thinking itself, is natural, conventions-free. According to this reasoning, which follows McNeill's, iconic gestures that realize spatio-visual concepts are transparent and do not obey social conventions. The presumed transparency of iconic signs has been a subject of a wide-ranging debate, and Eco (1976, 1984) and Kress and Van Leeuwen (2006), among others, have argued convincingly that the presumption is difficult to sustain. Furthermore, Lakoff and Johnson (1982) themselves maintain that conventions are involved even in the perceptions and motor actions from which their concepts, literal or metaphorical derive.

As well as testing the conventionality or lack thereof of the 'presenting' gesture which, according to McNeill's theory, should not be subject to standards of form, Parrill also tested for comparison the conventionality of a gesture that is widely assumed to be conventional, viz. the 'Okay' sign, usually performed by a circle made by a joined thumb and forefinger, with the rest of the fingers extended, facing upwards.

The results of Parrill's study did not confirm McNeill's theory since the 'Okay' gesture was not judged to be more conventionalized than the 'presenting' gesture. Parrill tries to justify the negative results in various ways by reference to the experimental situation, of which the most plausible seems that because emblems may be further below the level of consciousness than language, more directly probing questions should have been used. The main issue, however, is not why the modified versions of the 'Okay' signs were not judged to be unconventional, but why the different versions of the 'presenting' sign were judged as if they were conventional on some measures. According to McNeill's theory, there are meant to be no conventions involved in the formation of co-verbal gestures and, whatever the attempts to explain the results by recourse to some hypothetical reasons, the results of the experiment do not seem to confirm the theory.

Muller

Muller's chapter continues the same overall theme, explored in Parrill's chapter, of co-verbal gestures being instantial and improvised, which gives them the power to directly represent the spatio-visual thinking that tends to characterize metaphors. She however also focuses on another aspect of co-verbal metaphoric gestures, which is their 'dynamism' (see Nunez's chapter above). She thus prefers to talk about 'metaphoricity' rather than metaphor, emphasizing that one and the same metaphor may extend over a stretch of discourse and be repeatedly realized in similar gestures. Her concept of metaphoricity is obviously closely related to McNeill's catchment.

According to Muller, it is not just the instantial, non-conventionalized character of co-verbal gestures that provides evidence of the vitality of the metaphor but so does their dynamic quality. This dynamic quality is said to be due to metaphors, or metaphoricity, being a general cognitive process rather than belonging only to language. The evidence for this is precisely that the source domains of metaphorical concepts are realized in gestures as well as speech, and in other modes as well such as, for example, images (Forceville, 1996).

Muller follows McNeill's (e.g. 1992, 2005) opposition of cognition being processual, and so dynamic, versus language being product-like, and so static. Since Muller is able to identify repeated realizations of the same metaphor in gesture, she argues that this is evidence for metaphors being dynamic and therefore instantial. She argues that such gestures realize the source concepts 'directly', i.e. bypassing any convention.

As I already mentioned in relation to McNeill's chapter, I find it difficult to accept such a direct realization relationship. I am in agreement with Quinn, who

in her comment in the second part of the book ‘sharply disagrees’ with such metaphorical expressions being idiosyncratic and instantial, one-off, ‘online’ creations. In her set of data, which is similar to Muller’s, interviewees used metaphors in speech that were not strongly conventionalized (i.e. they were unlike ‘love is a journey’ and similar – see Lakoff and Johnson, 1982) and yet, they were used regularly by most of her subjects. Even in these cases, there does seem to be some ‘cultural guidance’ (Quinn: 253) that the speakers are subject to.

A comment on McNeill’s theory of co-verbal gestures

In what follows, I will critically discuss the main aspect of McNeill’s theory of co-verbal gestures: their purportedly non-systematic character. Together with their presumed context-sensitive and idiosyncratic nature, this, according to McNeill, makes them a communication mode that is different from, but complementary to, language. These three aspects of McNeill’s theory are the most important and the non-systematic aspect is more crucial than the other two, which may be said to be its consequences.

According to McNeill (e.g. 1992, 2005), for gestures to be systematic, there would have to be listable gesture morphemes with precise forms and regular correspondences between forms and meanings fixed by a social convention. McNeill (1992: 36–37) acknowledges that there is no conscious agreement made by the speakers of a language to observe its rules. However, it is true that the speakers do generally recognize when something goes wrong, and how easily they recognize these missteps likely depends on how conscious they are of the relevant normal patterns. One could probably argue that, in lexicogrammar, the closer to the formal, perceivable surface of the grammatical form the patterns are, the less abstract they are, and the more likely the speakers will be aware of them and of possible departures from them. At the same time, the closer the patterns or items are to clearly realizing some perceivable referent, the less abstract and more accessible to consciousness they likely are, too.

Halliday (1987) proposes a scale of consciousness of lexicogrammatical patterns, according to which lexical items are the most conscious and transitivity patterns the least, with group and phrase classes, and derivational morphology, in between. This makes sense since lexical items are completely realized in the form, group and phrase classes and derivational morphemes have quite obvious formal realizations – they are what Whorf (1958) called ‘phenotypes’, while the realization of transitivity patterns in forms is indirect and much less obvious, which is why Halliday (e.g. 1987), following Whorf, has called them cryptotypes. Somewhere in the middle of the above scale of consciousness could be added grammatical items such as pronouns, and general lexical items such as *somebody*, *something*, etc., since the referent of neither is clearly specified.

Co-verbal gestures are generally produced and received unconsciously, which contrasts with speech that most of the time is produced and received consciously: we know when we are saying something and when we are listening, but we generally do not notice when we are gesticulating or when others are. As well, speakers of all languages are taught linguistic conventions – first by their caretakers, and in literate cultures often much more explicitly in schools, and at times even in the workplace. All this brings linguistic patterns closer to consciousness, and nothing like that usually happens with co-verbal gestures.

This does not mean, though, that there are no regular gesture patterns produced repeatedly and systematically by different speakers. McNeill's insistence on comparing co-verbal gestures to morphemes is far too stringent and also rather limited. According to McNeill, if there were conventional and listable gesture morphemes, this would make gestures systematic like language, but since there are not, it means they are different, and this difference is at the basis of arguably the main gestures-language opposition that underpins his theory. Furthermore, although McNeill talks about morphemes, he only ever seems to compare gestures to lexical items (e.g. 1992: 41; 2005: 8). There is, of course, more to language than lexical items, or morphemes in general, such as the above-mentioned, more abstract transitivity patterns, with their indirect, cryptotypic realizations.¹

If one takes this more abstract approach to the study of gestures, and investigates how they realize transitivity patterns, one finds that, at this level, they are just as systematic as language (Martinec 2000c, 2004). Because of their abstract nature, as well as due to the gestures' unconscious production and reception, the transitivity patterns in gestures are quite surely deep below the level of the interactants' consciousness. But just like in language, even though they are deep below the level of consciousness, the transitivity patterns in gestures and their realizations in forms have regular correlations. They are thus systematic and likely conventionalized. Their production and interpretation simply obey conventions that are deeply unconscious.

Let us now look at McNeill's characterization of gestures as non-systematic in more detail. As was said above, he considers gestures being non-systematic to be an advantage since this way they are freed from the constraints of form, and can express whatever is relevant in the immediate, instantial context. He argues that this is because, while linguistic units' forms and meanings are prearranged and retrieved by the speaker, gestural forms are created by him/her on the spur of the moment, realizing the most relevant meanings at that point in the discourse. While linguistic units are categories, gestures thus typically appear to only be instantial, particular to a specific context. So what are the gestural forms determined by if not by the requirements of conventions?

According to McNeill (e.g. 2005: 48), they are determined by their meanings, and the meanings, in turn, are determined by the referents in the represented instancial context. This, according to McNeill, contrasts with linguistic morphemes, whose forms are arbitrarily related to meanings, and the meanings to their referents. It is because of these two lines of arbitrariness that conventions are needed to fix the relationships between the forms and meanings and the meanings and referents. Since the two relationships are iconic, or 'direct', in gestures, no such conventions are needed.

It has already been mentioned that the 'directness' of iconic representation has been vigorously contested. McNeill (2005: 58) does discuss Goodman's (1968) critique of photorealism, which he considers gestures to be exempt from, since they are not photorealistic representations. This leads him to argue that, furthermore, gestures are not representations but materializations of imagery, which in turn leads to his affirming (2005: 98) that there is no distinction between mental imagery (gestures' meanings) and gestures themselves. He draws on Fox's (1995) interpretation of Heidegger's concept of thought as cognitive being and on Merleau-Ponty's (1962) notion of meaning 'inhabiting' speech, to argue that, at least for the producer, gestures are a cognitive house of being in which the speaker dwells and which the receiver is invited to co-inhabit. Because of its rather nebulous nature, these concepts remind one of certain kinds of post-modern writing, and make it difficult to argue either for or against them with any rigor. This ultimately leads to such apparently incomprehensible statements, as in McNeill's chapter in the reviewed book (p. 168) that, since the metaphorical meanings in conventionally metaphorical gestures are already conventionalized, the conventionalization of the gesture forms and their relationship with the meanings is irrelevant and can be 'shaved off' by the principle of Occam's razor.

Be it as it may, McNeill takes the determination of gestural form by meaning, and the concomitant bypassing of conventions, characteristic of naïve iconism (cf also Buissac's comment about naturalistic fallacy in the second part of the reviewed book), to the extreme of erasing the meaning-form relationship altogether.

Despite the above reservations, let us assume that McNeill's idea of gestures as a mode of cognitive being from the perspective of the encoder has some validity. There is, however, still the decoder's perspective, from which gestures are considered to be representational signs, and McNeill seems to be writing mostly from this perspective since references to gestures' forms being driven by their meanings, as opposed to those of language depending on conventions, are densely interspersed throughout both McNeill (1992, 2005), and form the basis of many of his comparisons of the two modes. Gestures as representational signs, even though not photorealistic, are however

still subject to a broader critique of iconism to be found in, for example, Eco (1976), who in fact includes gestures as well as other non-photorealistic signs as examples. According to Eco (1976), arbitrariness of a relationship between form and meaning is not a prerequisite for the relationship to be conventional. So-called iconic signs are just as subject to conventions as arbitrary ones, since they, too, are instances of categories, and categories are fundamentally social.

As for the categories of co-verbal gestures, Martinec (2004) argues for their existence in both semantics and form. In the semantics, he presents classes of process: [action] or [state]; if [action], then [active] or [passive], etc., and of circumstance: [location] or [no location]. The formal units, or ranks, are forearm, hand and fingers. The classes at the forearm rank are: [movement] or [hold]; if [movement], then [force] or [no force], etc. The classes at the hand rank are: (when movement) [whole hand in the directions of movement] or [hand edge in the direction of movement]; if [hand in the direction of movement], then [vertical to the ground] or [horizontal to the ground], and so on. The classes at the fingers rank are: [bent], [in-between], [extended]. These classes are in principle the same as classes of the clause: [major] or [minor], etc., group or phrase: [nominal] or [verbal], etc. and word: [nominals] or [verbals], etc. The formal categories are of course proper to gestures, which is only correct, since applying formal linguistic categories to other semiotics has not proven to lead very far (see Kress and van Leeuwen, 2006). But they are just as generalized as those in language and their correlations with the semantic categories are just as systematic.

The suggestion of a presence of deeply unconscious conventions governing regular and generalized correspondences between the semantic and formal categories in co-verbal gestures is strengthened by the correspondences being far from direct, or transparent. Action processes, for example, such as walking or kicking a ball, are realized by a movement of the forearm, whereas state processes, such as a goal-keeper being at the far post of the goal, or one's being a sad human being, are realized by a lack of forearm movement. The iconicity of such realizations cannot exactly be said to be transparent – there is quite a difference between a person's kicking a ball and a forearm movement, just as there is one between a person standing next to a goal post and a forearm being still. And this is true of other such regular, systematic realizations – for example, that of the Goal of an action process by a hand turned in the direction of the movement and bent fingers (see Martinec, 2004).

This of course does not mean that gestures are not different from language. Apart from the more obvious differences such as gestures' not having a syntax and the consequences of this, the main less obvious difference appears to be that gestures, unlike language, are an indexical system (Martinec, 2004). This

means that they have to get their full meaning from the context they represent. In order to explain what this involves, two relationships from systemic-functional linguistics will be referred to: delicacy and instantiation.² Delicacy is a scale and the dimension that underlies it is that of general to specific. In lexicogrammar, types of process with their transitivity structures are said to be related in delicacy to lexical items, with the types of process at the most general end and lexical items at the most specific. This means that transitivity structures of the most general kind (e.g. Actor-Process:material-Goal) are described, or specified, by the smallest number of features, or dimensions of meaning, whereas lexical items, such as 'kick' are described by the greatest number, with more specific process types and more general lexical items in between. Instantiation relates the linguistic system and text – a text is an instance of the linguistic system, and similar textual instances make up text types, where the relationship between the instance and type is of abstraction. Generally speaking, it would seem that, since transitivity structures and more general lexical items are more abstract, they characterize text types, whereas the most delicate lexical items, being more concrete, characterize single textual instances. Text types, at the same time, realize situation types, while textual instances realize instantial situations (e.g. Gregory, 1988).

The types of process in gestures mapped out in Martinec (2000c, 2004) are less delicate and more abstract, and it seems that this is how many gestures are. Others, like some of those that McNeill (e.g. 1992: 106–108) focuses on when discussing different speakers' representations of a cartoon character climbing up a drain pipe, could perhaps be interpreted as being more delicate and less abstract.³ McNeill argues that, since different speakers represent a particular event in a narrative stimulus (a Disney cartoon) in their gestures differently, this means that gestures are idiosyncratic. The event in question is Sylvester the cat climbing up a drainpipe to get to Tweety the bird who is at the top of it. The source of variation between the different gestures however lies in the particulars of what the speakers focus on in the represented event, i.e. the different manners of climbing. It is quite likely that if McNeill had more subjects than just the five he did, some of them would focus on selecting the same aspect of Sylvester's climbing experience. In that case, it is likely that these same aspects would be represented by the same gestures. This seems to be the point that Quinn raises in her comment in the second part of the reviewed book. Even such more specific and less abstract gestures would thus form a class, or category, which would be instantiated in each of the subjects' multimodal text. The gestures would thus not be particular to each individual but a whole group, and some kind of 'cultural guidance' (Quinn), or unconscious convention (see above), would govern their forms and form-meaning correlations.

A similar critique applies to McNeill's (2005: 7–8) discussion of what could be interpreted as less abstract and more delicately specified gestures representing two different actions of bending. McNeill writes that the actions of bending a sheet of paper and bending a tree branch are represented by the same lexical item in language, but would be represented by very different gestural forms. This relates to his characterization of gestures as being context-sensitive, which is contrasted with the 'decontextualized' meaning of lexical items in language. However, there is no reason why more delicate and less abstract gestures should behave like language. It is in any case probable that any speaker who would use gestures to represent the actions of bending a tree branch and a sheet of paper would represent them by the same gesture form. They would also most likely use the same form to represent bending objects similar to a tree branch or to a sheet of paper. The gesture forms would then again form fairly delicate and concrete classes or categories, which would be instantiated in different textual instances. The decontextualized meaning of lexical items in language appears to be something of a myth in any case. Most lexical items in a dictionary have more than one meaning that are often defined in relation to different contexts, and this principle is reflected in, for example, Fillmore's (e.g. 1977, 1982) lexical semantics.

It does appear true, however, that the specification of the more delicate and less abstract gestures stops earlier than the specification of lexical items in language. Furthermore, as was said above, co-verbal gestures are generally produced and received well below the level of consciousness of lexical items in language. This means that the conventions that govern their forms and form-meaning correlations are rather less precise than those that concern lexical items in language – they are more like those that govern transitivity structures in language. These two characteristics are behind the full meaning of the gestures having to be retrieved from the represented context and thus behind their indexical quality. Without knowing such context, the meaning of a gesture for someone whose awareness might be drawn to it by an analyst is not 'The player kicked the ball' but 'someone did/does something'.

Apart from gestures being an indexical system, another difference in relation to language is their iconic realization relationship between expression and form as opposed to the arbitrary one in language.⁴ The levels of expression and form in gestures are thus collapsed into one, since there is only one level of agnation instead of two. It is this that leads to a greater iconicity between gestural meanings and forms and also to the 'processual', or 'analogue', nature of the realization relationship.⁵ It is this kind of relationship that allows gestures to represent the actions that form the source of the 'live' metaphors that so much emphasis is put on in the reviewed book. This is

of course true of the meanings of non-metaphorical gestures as well. The iconic relationship between the gesture forms and meanings simply enables gestures to better represent spatially based meanings than language. Such meanings could often also be represented in language, however, they would have to be described since there are no lexical items that would directly fit them.

A good example is McNeill's (1992: 253–256) detailed analysis of a growth point that contains the concept of 'buying' and of the contribution of gestures and language to its unfolding in the following stretch of multimodal text. The concept of 'buying' is part of Fillmore and Atkins' (1991) frame (similar to a situation type) of commercial transaction. McNeill maintains that, apart from the linguistic concept 'buy', the growth point is also made up of a schematized image of a buying encounter, and this contains the movement of handing over the goods. Furthermore, the movement may have a slanting quality that may reflect the different relative positions of the seller and buyer, either literal or metaphorical (in terms of their relative status). The verbal aspect of the growth point is realized in the linguistic structures and lexical items that make up the following stretch of text. This is also true of the handing-over movement, except for the slanting aspect of it, which is realized in gesture. The verbal concept of buying and the handing-over movement belong to the frame, which is at the level of a situation type. The slanting aspect of the movement may be said to belong to a particular, instancial situation, and is not part of the typical linguistic realizations. It is such selections from the represented context, particular to its instancial qualities, that are realized by gestures, according to McNeill, and it is these more delicate and instancial kinds of meaning that the more delicate and concrete gestural items tend to realize.

It does appear true that, although the slanting quality of the movement could be described in language, it normally would not be. A novelist like Tolstoy, whose style includes detailed descriptions of perceptions, may perhaps focus on expressing something like that for aesthetic reasons, but verbal art is a departure from everyday norm. It is thus without a doubt that a gesture serves the purpose of representing such a spatially based, more delicate and instancial aspect of the situation better than language. And this is related to the descriptions of gestures realizing the live aspects of metaphors in the relevant chapters of the volume under review. It is however also true that one cannot represent an aspect like the slanting quality of a movement without representing the movement itself at the same time. And this less delicate and more abstract movement would be represented in, apart from the linguistic structures and items that McNeill talks about, a gestural transitivity structure-form correlation of the kind presented in Martinec (2000c, 2004).

Notes

1. McNeill follows a no longer dominant, formal linguistic model (see e.g. McNeill: 71–73) based on the classic Saussurean dichotomies. This being the case, it is of course not surprising that he finds so many differences between gestures and language.
2. The combination of delicacy and instantiation was already present in Martinec (2004), although the main emphasis was on delicacy. I explicitly interpreted McNeill's more delicate gestures' meanings as instantial and the meanings of my less delicate indexes as typical, both having to be fully retrieved from the represented context. Indexicality in Martinec (2004) is thus not a 'common sense' concept because both delicacy and instantiation are used to explain it.
3. An alternative interpretation would be to analyze the same gestures as less delicate and more abstract process types combined with circumstances or 'aspects' (Martinec, 2000a). The various ways of the character's climbing up the drainpipe would thus be interpreted as material processes combined with aspects of Manner.
4. 'Iconic' is here used in a qualified sense that takes into account the critiques of these concepts mentioned earlier.
5. The presumed arbitrariness of the relationships between meanings and forms, and meanings and referents, which McNeill contrasts with the iconicity of these relationships in gestures has been challenged, among others, by Halliday (e.g. 1994), Hasan (1985) and in much of the work on cognitive semantics and grammar (e.g. Lakoff, 1987; Langacker, 2000). The iconicity of the relationships in language does not, however, appear to reach the level of iconicity in gestures.

References

- Calbris, G. (1990) *The Semiotics of French Gestures*. Bloomington, IN: University of Indiana Press.
- Eco, U. (1976) *A Theory of Semiotics*. Bloomington, IN: Indiana University Press.
- Eco, U. (1984) *Semiotics and the Philosophy of Language*. Bloomington, IN: University of Indiana Press.
- Ekman, P. and Friesen, W. V. (1969) The repertoire of non-verbal behaviour: Categories, origins, usage and coding. *Semiotica* 1 (1): 49–98.
- Fauconnier, G. and Turner, M. (2002) *The Way We Think: Conceptual Blending and the Mind's Hidden Complexities*. New York: Basic Books.
- Forceville, C. (1996) *Pictorial Metaphor in Advertising*. London: Routledge. doi:10.4324/9780203272305
- Fillmore, C. J. (1977) Topics in lexical semantics. In R. W. Cole (ed.) *Current Issues in Linguistic Theory*, 76–138. Bloomington, IN: Indiana University Press.
- Fillmore, C. J. (1982) Frame semantics. In Linguistic Society of Korea (eds) *Linguistics in the Morning Calm*, 111–137. Seoul: Hansin Publishing Co.
- Fillmore, C. J. and Atkins, B. T. (1991) Towards a frame-based lexicon: The semantics of RISK and its neighbors. In A. Lehrer and E. Kittay (eds) *Frames and Fields*. New York: Erlbaum.

- Fox, B. (1995) On the embodied nature of grammar: Embodied being-in-the-world. Plenary talk at the International Conference on Functional Grammar. University of New Mexico.
- Goodman, N. (1968) *Languages of Art: An Approach to a Theory of Symbols*. Indianapolis, IN: Bobbs-Merrill.
- Gregory, M. (1985a) Towards communication linguistics: A framework. In J. D. Benson and W. S. Greaves (eds) *Systemic Perspectives on Discourse*, Vol 1, 000–000. Norwood, MA: Ablex.
- Gregory, M. (1985b) Phasal analysis within communication linguistics: Two contrasting discourses. In J. Copeland (ed.) *Proceedings of the Second Rice University Symposium on Linguistics and Semantics*. Chicago, IL: University of Chicago Press.
- Gregory, M. (1988) Generic situation and register: a functional view of communication. In J. D. Benson, M. J. Cummings and W. S. Greaves (eds) *Linguistics in a Systemic Perspective*, 303–000. Amsterdam: Benjamins.
- Halliday, M. A. K. (1987) Language and the order of nature. In N. Fabb, D. Attridge, A. Durant and C. MacCabe (eds) *The Linguistics of Writing*, 000–000. Manchester: Manchester University Press.
- Halliday, M. A. K. (1994) *An Introduction to Functional Grammar*, 2nd edition. London: Arnold.
- Hasan, R. (1985) Meaning, text and context: Fifty years after Malinowski. In J. D. Benson and W. S. Greaves (eds) *Systemic Perspectives on Discourse*, Vol I, 000–000. Norwood, MA: Ablex.
- Kendon, A. (1972) Some relationships between motion and speech. In A. Siegman and B. Pope (eds) *Studies in Dyadic Communication*, 000–000. New York: Pergamon.
- Kendon, A. (1980) Gesticulation and speech: two aspects of the process of utterance. In M. R. Key (ed) *The Relationship of Verbal and Nonverbal Communication*, 207–228. The Hague: Mouton.
- Kress, G. and Van Leeuwen, T. (2006) *Reading Images: The Grammar of Visual Design*, 3rd edn. London: Routledge.
- Lakoff, G. (1987) *Women, Fire, and Dangerous Things: What Categories Reveal about Thought*. Chicago, IL: The University of Chicago Press.
- Lakoff, G. and Johnson, M. (1980) *Metaphors We Live By*. Chicago, IL: University of Chicago Press.
- Lakoff, G. and Nunez, R. (2000) *Where Mathematics Comes From: How the Embodied Mind Brings Mathematics into Being*. New York: Basic Books.
- Langacker, R. W. (2000) *Grammar and Conceptualization*. Berlin: Mouton.
- Martinec, R. (1998) Cohesion in action. *Semiotica*, 120 (1/2): 161–180. doi:10.1515/semi.1998.120.1-2.161
- Martinec, R. (2000a) Construction of identity in M. Jackson's *Jam*. *Social Semiotics*, 10 (3): 313–329. doi:10.1080/10350330050136370

- Martinec, R. (2000b) Rhythm in multimodal texts. *Leonardo*, 33 (4): 289–297. doi:10.1162/002409400552676
- Martinec, R. (2000c) Gestures used in speaking as a resource and as spontaneous creations. Plenary talk at the 27th Linguistics Association of Canada and the United States Forum, Rice University, Houston, TX.
- Martinec, R. (2002) Rhythmic hierarchy in monologue and dialogue. *Functions of Language*, 9 (1): 39–59.
- Martinec, R. (2004) Gestures which co-occur with speech as a systematic resource: The realization of experiential meanings in indexes. *Social Semiotics*, 14 (2): 193–213. doi:10.1080/1035033042000238259
- Martinec, R. and A. Salway (2005) A system for image-text relations in new (and old) media. *Visual Communication*, 4 (3): 337–371. doi:10.1177/1470357205055928
- 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.
- Merleau-Ponty, M. (1962) *Phenomenology of Perception*. London: Routledge.
- Peirce, C. S. (1955) Logic as semiotic. In J. Bucher (ed) *Philosophical Writings of Peirce*, 98–119. New York: Dover.
- Tuite, K. (1993) The production of gesture. *Semiotica*, 93 (1/2): 83–105. doi:10.1515/semi.1993.93.1-2.83
- Vygotsky, L. S. (1987) *Thought and Language*. Cambridge, MA: MIT Press.
- Whorf, B. J. (1956) *Language, Thought and Reality*, J. B. Carroll (ed). Cambridge, MA: MIT Press.

Book reviewed

- Cienki, A. and C. Muller (eds) (2008) *Gesture and Metaphor* Amsterdam/Philadelphia, PA: Benjamins.