Chapter 32

Gesture

Andy Lücking

Goethe-Universität Frankfurt

The received view in (psycho-)linguistics, dialogue theory and gesture studies is that co-verbal gestures, i.e. hand and arm movement, are part of the utterance and contribute to meaning. The relationships between gesture and speech obey regularities which are consequently captured in grammar by means of a gesture-grammar interface. This chapter provides basic snapshots from gesture research, reviews constraints on speech-gesture integration and summarises their implementations into HPSG frameworks. Pointers to future developments conclude the exposition. Since there are already a couple of overviews on gesture such as Özyürek (2012); Wagner et al. (2014); Abner et al. (2015), this chapter aims at distinguishing itself in that it provides a guided tour, where the guiding force is drawn from semantics and general semiotics and focuses, of course, on gesture work within HPSG.

1 Why gestures?

People talk with their whole body. A verbal utterance is couched into an intonation pattern that, *via* prosody, articulation speed or stress, function as *paralinguistic* signals (e.g. Birdwhistell 1970). The temporal dimension of paralinguistics gives rise to *chronemic* codes (Poyatos 1975; Bruneau 1980). *Facial expressions* are commonly used to signal emotional states (Ekman & Friesen 1978), even without speech (Argyle 1975), and are correlated to different illocutions of the speech

acts performed by a speaker (Domaneschi et al. 2017). Interlocutors use *gaze* as a mechanism to achieve joint attention (Argyle & Cook 1976) or provide social signals (Kendon 1967). Distance and relative direction of speakers and addressees are organised according to culture-specific radii into social spaces (*proxemics*, Hall 1968). Within the inner radius of private space, tactile codes of *tacesics* (Kauffman 1971) are at work. Since the verbal and nonverbal communication means of face to face interaction may occur simultaneously, *synchrony* is a signal in its own (Wiltshire 2007). A special chronemic case in signalling at the right moment – or, for that matter, to the contrary: missing the right moment (an aspect of communication dubbed *kairemics* by Lücking & Pfeiffer (2012: 600)). Besides the manifold areas of language use, the conventionalised, symbolic nature of language secures language's primacy in communication, however (de Ruiter 2004). For thorough introductions into semiotics and multimodal communication see Nöth (1990), Posner et al. (1997–2004) or Müller et al. (2013–2014).

The most conspicuous non-verbal communication means of everyday interaction are hand and arm movements, known as *gestures*. In seminal works, McNeill (1985; 1992) and Kendon (1980; 2004) argue that co-verbal gestures, i.e. hand and arm movements, can be likened to words in the sense that they are part of a speaker's utterance. Accordingly, integrated speech-gesture production models have been devised (Kita & Özyürek 2003; de Ruiter 2000; Krauss et al. 2000) that treat utterance production as a multimodal process. Given gestures' imagistic and often spontaneous character, it is appealing to think of them as "postcards from the mind" (de Ruiter 2007: p. 21). Clearly, given this entrenchment in speaking, non-verbal communication means have repercussions to areas hitherto taken to be purely linguistic (in the sense of being related to the verbal domain). This section highlights some phenomena particularly important for grammar, including, for instance, *mixed syntax* (Slama-Cazacu 1976), or *pro-speech gesture*:

(1)   He is a bit [*circular movement of index finger in front of temple*].

In (1), a gesture replaces a position that is usually filled by a syntactic constituent. The gesture is emblematically related to the property of *being mad* so that the mixed utterance from (1) is equivalent to the proposition that the referent of *he* is a bit mad.

Probably the most conspicuous gestures accompanying speech are pointing gestures (or deictic gestures). Given that the pronoun in (1) is used exophorically (that is, referring to someone from the physical utterance situation instead to a discourse referent from the preceding discourse), a pointing signal is required indicating where to find the referent. With regard to deictic gestures Fricke (2012:

Sec. 5.4) argues that deictic words within noun phrases – her prime example is German *so* (*like this*) –, provide a *structural*, that is, *language-systematic* integration point between the vocal plane of conventionalized words and the non-vocal plane of body movement. Therefore, on this conception, not only utterance production but *grammar* is inherently multimodal.

Moving from sentence to dialogue, *interactive gestures* are bound up with , among others (Bavelas et al. 1992; 1995). For instance, pointing gestures can be used to indicate the next speaker (Rieser & Poesio 2009). Thus, inasmuch as grammar is multimodal, also dialogue theory has to deal with specific non-verbal interaction means (see also Lücking, Ginzburg & Cooper (2019), Chapter 30 of this volume, this volume).

It should be noted, however, that not all gestures are as clearly a part of a sentence or an utterances as those in the examples given so far. A prominent counter-example in particular for interactive gesture views is gesturing on the telephone (see Bavelas et al. (2008) for an overview of a number of respective studies). Since such gestures are not observable for the addressee, they cannot reasonably taken to be a constituent of the intended content. Rather, "telephone gestures" seem to be purely speaker-oriented, presumably facilitating word retrieval. Furthermore, the lion's share of everyday gestures seem to consist of rather sloppy movements that do not contribute to the content of the utterance – they are *contingent* as opposed to be an obligatory semantic component (Lücking 2013). Based on such dividing evidence, gestures have broadly been distinguished into *foreground* vs. *background* gestures (Cooperrider 2017). This distinction aims to cover a basic difference in gestures and in the way a gesture is integrated into an utterance. Roughly, foreground gesture, but not background gestures, are semantically interpretable components of an utterance. This chapter focuses on foreground gestures.

2 Kinds of gestures

Gestures come in various kinds and are usually assigned to a number of gesture classes. These gesture classes are characterized by the function performed by a gesture and the meaning relation the gesture bears to its content (if any). A classic taxonomy consists of the following classes (McNeill 1992):

- iconic (or representational) gestures. Spontaneous hand and arm movements that are commonly said to be based on some kind of resemblance

relation.[1]. Iconic gestures employ a mode of representation such as *drawing*, *modeling*, *shaping* or *placing* (Streeck 2008; Müller 1998).

- deictic gestures (pointing). Typically hand and arm movements that perform a demonstration act. In which way pointing is standardly accomplished is subject to culture-specific conventions, though (Wilkins 2003). In principle, any extended body part, artifact or locomotor momentum will serve the demonstrative purpose. Accordingly, there are deictic systems that involve lip-pointing (Enfield 2001) or nose-pointing (Cooperrider & Núñez 2012). Furthermore, under certain circumstances, pointing with the eyes (gaze-pointing) is also possible (Hadjikhani et al. 2008). Note further that the various deictic means can be interrelated. For instance, manual pointing can be differentiated by cues of head and gaze (Butterworth & Itakura 2000).

- beats (rhythmic gestures). Hand and arm movements that are coupled to the intonational or rhythmic contour of the accompanying speech. Beats lack representational content but are usually used to an emphasising effect.

- emblem (lexicalized gestures). In contrast to the other classes, emblems are special in that they follow a fully conventionalized form-meaning relation. A common example in Western countries is the thumbs-up gesture, signalling "approval or encouragement" (Merriam Webster online dictionary[2]).

In addition to classifying gestures according to the above-given functional groups, a further distinction is usually made with regard to the ontological place of their referent: representational and deictic gestures can relate to concrete or to abstract objects or scenes. For instance, an iconic drawing gesture can metaphorically display the notion "genre" (McNeill 1992: 14). Likewise, pointing into a location within gesture space as a proxy for a real-world place is a point in case. An example is provided in Figure 5 below. Accordingly, gestures can be cross-classified into concrete and *abstract* or *metaphorical* ones.

Besides categorially distinguished gesture classes, gestures can also be specified according to multi-dimensional or functional characteristics (McNeill 2005;

---

[1]But see footnote 5 in Section 4.4 for pointers to critical discussions of resemblance as a sign-bearing relation.

[2]https://www.merriam-webster.com/dictionary/thumbs-up, accessed 20th August 2018. The fact that emblems can be lexicalized in dictionaries emphasizes their special, conventional status among gestures.

Gerwing & Bavelas 2004). See Özyürek (2012) for an overview of various gesture classification schemes.

On the most basic, kinematic level the movement of a prototypical gesture follows an "anatomic triple": gestures have to be partitioned into at least a preparation, a stroke, and a retraction phase (Kendon 1972). The stroke is the movement part that carries the gesture's meaning. Preparation and retraction phase bring hand and arms into respectively out of the stroke. Unless stated otherwise, when talking about gestures what follows, the stroke phase, which is the "gesture proper" or the "semantic interpretable" phase, is referred to.

In any case, the spontaneous, usually co-verbal hand and arm movement considered in this chapter are different from the signed signs of sign languages (see Steinbach & Holler (2019), Chapter 31 of this volume) and pantomime (not spontaneous and co-verbal).[3]

3 Basic empirical phenomena of grammatical gesture integration

With regard to grammar-gesture integration, three main phenomena have to be dealt with:

- What is the meaning of a gesture? How to assign semantic representations or truth conditions to hand and arm movements?

- What is the affiliate of a gesture, that is, its verbal attachment site?

- What is the result of multimodal integration, that is, composing verbal and non-verbal meaning?

The following example, (2) and Figure 1, (taken from Lücking (2013: 189)) illustrates these issues:

(2) Ich g[laube das   sollen TREP]pen   sein.
     I   think   those should STAIRCAses be.

     'I think those should be staircases.'

The first syllable of the German noun *Treppen* (*staircases*) carries main stress, indicated by capitalization. The square brackets indicate the temporal overlap

---

[3]In languages like German the difference between free gesticulation and sign language signs is also reflected terminologically: the former are called *Gesten*, the latter *Gebärden*.

Figure 1: "I think those should be staircases."

between speech and gesture stroke, which is shown in Figure 1. The gesture attributes to the noun it attaches to: from the multimodal utterance the observer retrieves the information that the speaker talks about spiral staircases. This interpretation assumes that the common noun is the affiliate of the iconic gesture. Obviously, mere temporal synchrony is too weak to be an indicator of affiliation. In fact, there are speech-gesture affiliations without temporal overlap between gesture and verbal affiliate at all (e.g. Lücking et al. 2004). Therefore, temporal overlap or vicinity is just one indicator of affiliation, a second one is intonation: a gesture is usually related to a stressed element in speech (McClave 1994; Nobe 2000; Loehr 2004; 2007). As a result, multimodal communication gives rise to complex "peak pattern" (Tuite 1993; Loehr 2004; Jannedy & Mendoza-Denton 2005).

The interpretation of a gesture changes with different affiliations. Suppose the gesture from Figure 1 is produced in company to stressed *glaube* (*think*) instead of *staircases*:

(3) Ich G[LAUbe das   sollen Trep]pen sein.
     I   THINK   those should staircases be

     'I think those should be stairscases.'

Now the spiral movement is interpreted as a metaphorical depiction of a psychological process. Thus, the interpretation of a gesture depends on the integration point (affiliation), which in turn is marked by prosody. Note that the resulting may express a richer content than speech alone, as in (2), or a content equivalent to speech content, as in (3); it can even express less than speech or contradict

speech:[4]

> The nonverbal act can repeat, augment, illustrate, accent, or contradict the words; it can anticipate, coincide with, substitute for or follow the verbal behavior; and it can be unrelated to the verbal behavior. (Ekman & Friesen 1969: 53)

Given the linguistic significance of gestures as sketched in the preceding sections grammar-oriented accounts on speech-gesture integration have recently been developed that try to deal with (at least one of) the three basic phenomena, though with different settings of priorities, including Alahverdzhieva (2013), Alahverdzhieva & Lascarides (2010), Ebert (2014), Giorgolo (2010), Giorgolo & Asudeh (2011), Lücking (2013; 2016), Rieser (2008; 2011; 2015), Rieser & Poesio (2009), Schlenker (2018).

The HPSG-related approaches are briefly reviewed below. For grammars for sign languages see Steinbach & Holler (2019), Chapter 31 of this volume. There also belong investigations into iconic features of American Sign Language (Schlenker et al. 2013; Schlenker 2014).

4 Gestures in HPSG

Integrating a gesture's (pointing) contribution to speech was initiated in computer science. Coincidentially, these early works used typed feature structure description akin to HPSG grammars. Though linguistically limited, the crucial invention has been a *multimodal chart parser*, that is, an extension of chart parsing that allows to process input on two modalities (namely speech and gesture). Such approaches are reviewed in Section 4.1. Afterwards, a more elaborate gesture representation format is introduced that allows to encode the observable form side of a gesture in terms of kinematically derived attribute-value structures (Section 4.2). Following the basic semiotic distinction between deictic (or indicating or pointing) gestures and iconic (or representational or imagistic) gestures, the analysis of both classes of gestures is exemplified in Sections 4.3 and 4.4, respectively.

---

[4]In case of contradiction, that, speech-gesture mismatch, the resulting multimodal utterance is perceived as ill-formed and induces N400 effects (Wu & Coulson 2005; Kelly et al. 2004).

4.1 Precursors

Using typed feature structure descriptions in order to represent the form and meaning of gestures goes back to computer science approaches to human-computer interaction (HCI). The *QuickSet* system (Cohen et al. 1997) allows users to operate on a map and move objects or create barb wires (the project was funded by a grant from the US army) by giving verbal commands and manually indicating coordinates. The system processes voice and pen (gesture) input by assigning signals from both media representations in the form of attribute-value matrices (AVMs) (Johnston 1998; Johnston et al. 1997). For instance, *QuickSet* will move a vehicle to a certain location on the map when asked to *Move this*[☞] *motorbike to here*[☞], where '☞' represents an occurrence of touch gesture (i.e., pen input).

Since a linguistic unification-based grammar rests on a conventional, "unimodal" parser, Johnston (1998) and Johnston et al. (1997) developed a *multimodal chart parser*, which are still a topic of computational linguistics (Alahverdzhieva et al. 2012) (see also Bender & Emerson (2019), Chapter 29 of this volume). A multimodal chart parser consists of two or more layers and allows for layer-crossing charts. The multimodal NP *this*[☞] *motorbike*, for instance, is processed in terms of a multimodal chart parser covering of a speech (s) and a gesture (g) layer:

(4)   s:

g:

A multimodal chart or *multichart* is defined in terms of sets of identifiers from both layers. Possible multicharts from (4) include the following ones:

(5)   multichart 1: {[s,0,1], [g,3,4]}
      multichart 2: {[s,1,2], [g,3,4]}
      ...

The basic rule for integrating spatial gestures with speech commands is the *basic integration scheme* (Johnston 1998; Johnston et al. 1997), reproduced in (6):

Draft of 1st April 2019, 11:14

(6)
$$
\begin{bmatrix}
\text{LHS} : & \begin{bmatrix}
\text{CAT} : & command \\
\text{MODALITY} : & \boxed{2} \\
\text{CONTENT} : & \boxed{1} \\
\text{TIME} : & \boxed{3}
\end{bmatrix} \\
\text{RHS} : & \begin{bmatrix}
\text{DTR1} : & \begin{bmatrix}
\text{CAT} : & located\_command \\
\text{MODALITY} : & \boxed{6} \\
\text{CONTENT} : & \boxed{1}[location\ \boxed{5}] \\
\text{TIME} : & \boxed{7}
\end{bmatrix} \\
\text{DTR2} : & \begin{bmatrix}
\text{CAT} : & spatial\_gesture \\
\text{CONTENT} : & \boxed{5} \\
\text{MODALITY} : & \boxed{9} \\
\text{TIME} : & \boxed{10}
\end{bmatrix}
\end{bmatrix} \\
\text{CONSTRAINTS} : & \left\{
\begin{array}{l}
\text{OVERLAP}(\boxed{7},\boxed{10}) \vee \text{FOLLOW}(\boxed{7},\boxed{10},4s) \\
\text{TOTAL-TIME}(\boxed{7},\boxed{10},\boxed{3}) \\
\text{ASSIGN-MODALITY}(\boxed{6},\boxed{9},\boxed{2})
\end{array}
\right\}
\end{bmatrix}
$$

The AVM in (6) implements a mother-daughter structure along the lines of a context free grammar rule, where a left-hand side (LHS) expands to a right-hand side (RHS). The right-hand side consists of two constituents (daughters DTR1 and DTR2), a verbal expression (*located_command*) and a gesture. The semantic integration between both modalities is achieved in terms of structure sharing, see tag $\boxed{5}$: the spatial gesture provides the location coordinate for the verbal command.

The bimodal integration is constrained by a set of restrictions, mainly regulating the temporal relationship between speech and gesture (see tags $\boxed{7}$ and $\boxed{10}$ in the CONSTRAINTS set): the gesture may overlap with its affiliated word in time, or follow it in at most four seconds.

An integration scheme akin to that displayed in (6) also underlies current grammar-oriented approaches to deictic and iconic gestures (see Sections 4.3 and 4.4 below).

4.2 Representing gestures with AVMs

Representing the formal features of gestures in terms of attribute-value matrices has been initiated in robotics (Kopp et al. 2004). A representation format that captures the "phonological", physical-kinematic properties of a gesture is designed according to the moveable junctions of arms and hands. For instance, the representation of the gesture in Figure 1 according to the format used in Lücking et al. (2010) is given in (7):

(7)

$$
\begin{bmatrix}
\text{right hand} & & \\
\text{HANDSHAPE} & \begin{bmatrix} \text{SHAPE} & G \\ \text{PATH} & 0 \\ \text{DIR} & 0 \end{bmatrix} \\
\text{PALM} & \begin{bmatrix} \text{ORIENT} & PAB{>}PAB/PUP{>}PAB \\ \text{PATH} & 0 \\ \text{DIR} & 0 \end{bmatrix} \\
\text{BOH} & \begin{bmatrix} \text{ORIENT} & BUP{>}BTB/BUP{>}BUP \\ \text{PATH} & arc{>}arc{>}arc \\ \text{DIR} & MR{>}MF{>}ML \end{bmatrix} \\
\text{WRIST} & \begin{bmatrix} \text{POSITION} & P\text{-}R \\ \text{PATH} & line \\ \text{DIR} & MU \\ \text{DIST} & D\text{-}EK \\ \text{EXTENT} & small \end{bmatrix} \\
\text{SYNC} & \begin{bmatrix} \text{CONFIG} & BHA \\ \text{REL.MOV} & LHH \end{bmatrix}
\end{bmatrix}
$$

The formal description of a gestural movement is given in terms of the handshape, the orientations of the palm and the back of the hand (BOH), the movement trajectory (if any) of the wrist and the relation between both hands (synchronicity, SYNC). The handshape is drawn from the fingerspelling alphabet of American Sign Language, as illustrated in Figure 2. The orientations of palm and back of hand are specified with reference to the speaker's body (e.g., *PAB* encodes "palm away from body" and *BUP* encodes "back of hand upwards"). Movement features are specified with respect to the wrist: the starting position is given and the performed trajectory is encoded in terms of the described path and the direction and extent of the movement. Position and extent are given with reference to the *gesture space*, that is the structured area within the speaker's immediate reach (McNeill 1992: 86–89) – see the left-hand side of Figure 3. Originally, McNeill considered the gesture space as "a shallow disk in front of the speaker, the bottom half flattened when the speaker is seated" (McNeill 1992: 86). However, acknowledging also the distance of the hand with respect to the speaker's body (feature DIST) turns the shallow disk into a volume, giving rise to the three-dimensional model displayed in the right-hand side of Figure 3. The gesture space regions known as *center-center*, *center* and *periphery*, possible modified by location modifiers (*upper right*, *right*, *lower right*, *upper left*, *left*, *lower left*), are now modeled as nested hexahedrons. Thus, gesture space is structured according to all three body axes: the sagittal, the longitudinal, and the transversal axis. Annotations straightfowardly transfer to the three-dimensional gesture space model. Such a
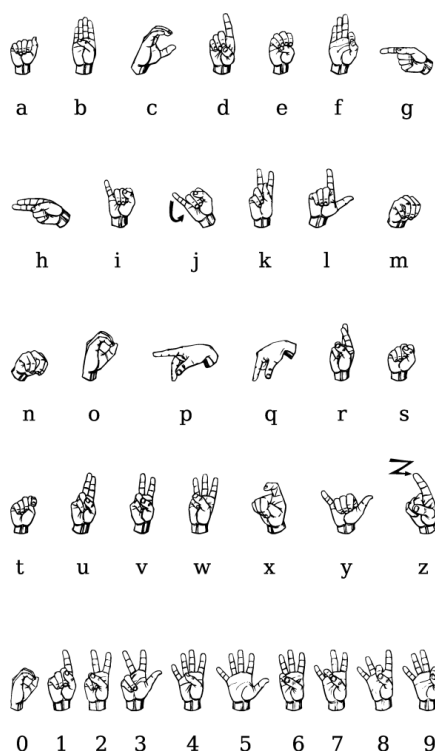
Figure 2: American Sign Language fingerspelling alphabet (image released to Public Domain by user Ds13 in the English Wikipedia at 18th December 2004, https://commons.wikimedia.org/wiki/File:Asl_alphabet_gallaudet.png).

three-dimensional gesture space model is assumed throughout this chapter. Rectangular and roundish movement are distinguished in terms of *line* or *arc* values of feature PATH. An example is given in Figure 4. For a documentation of gesture representation along this line see Lücking et al. (2016). For related formats see Martell (2002) and Gibbon et al. (2003). For a different approach see Lausberg & Sloetjes (2009).

The "phonological" representation of gestures is slightly refined in Lücking (2016: 385–386), where gestures are analysed in terms of a formal event framework (Fernando (2011), Cooper & Ginzburg (2015: Sec. 2.7)) in order to account for Gestalt properties of gesture movements (see also the example discussed by Rieser (2008)).

Figure 3: Gesture Space (left hand side is simplified from McNeill (1992: 89).) Although originally conceived as a structured "shallow disk" McNeill (1992: 86), adding distance information gives rise to a three-dimensional gesture space model as illustrated on the right-hand side and assumed in the main text.
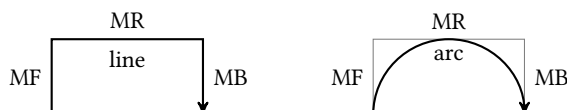

Figure 4: The same sequence of direction labels can give rise to an open rectangle or a semicircle, depending on the type of concatenation (Lücking 2016: 385).

4.3 Pointing Gestures

Pointing gestures are *the* referring device: they probably pave a way to reference in both an evolutionary and language acquisition perspective. (Bruner 1998; Masataka 2003; Matthews et al. 2012); they are predominant inhabitants of the "deictic level" of language, interleaving the symbolic (and the iconic) ones (Levinson (2008), see also Bühler (1934)); they underlie reference in *Naming Games* in computer simulation approaches (Steels 1995) (for a semantic assessment of naming and categorisation games see Lücking & Mehler (2012)). The referential import of pointing gesture has been studied experimentally to some detail (Bangerter & Oppenheimer 2006; Kranstedt et al. 2006b,a; van der Sluis & Krahmer 2007). As a result, it turns out that pointings do not rest on a direct "laser"- or "beam" mechanism (McGinn 1981). Rather, they serve a (more or less rough) locating function (Clark 1996) that can be modelled in terms of a *pointing cone* (Lücking et al. 2015). These works provide an answer to the first basic question

Figure 5: Driving towards/heading straight

(cf. Section 3): pointing gestures have a "spatial meaning" which focuses or highlights a region in relation to the direction of the pointing device. Such a spatial semantic model has been introduced in Rieser (2004) under the name of *region pointing*, where the gesture adds a locational constraint to the restrictor of a noun phrase. The spatial model is also adopted in Lascarides & Stone (2009), where the region denoted by a pointing is represented by a vector $\vec{p}$. This region is an argument to function v, however, which maps the projected cone region to $v(\vec{p})$, the space-time talked about, which may be different from the gesture space (many more puzzles of local deixis are collected by Klein (1978) and Fricke (2007)).

Let us illustrate some aspects of pointing gesture integration by means of a real world example (taken from dialogue V5 of the SaGA corpus (Lücking et al. 2010)). The speaker in (8) and Figure 5 places a pond with his left hand into gesture space. He then points at his left hand/the pond (on *dual points* see Goodwin (2003), on *deferred reference* see Quine (1950); Nunberg (1993)), depicting *driving towards* or *heading straight*. The movement aspect is subsequently enacted by moving the right hand/driver towards the left hand/pond.

(8)   wenn du dort eingefahren bist, fährst Du geradeaus auf einen Teich zu ..
      'n Teich .. und .. [*gesture from Figure 5 occurs here*]

      when you drive in there , you're heading straight for a pond .. a pond .. and ..
      [*gesture from Figure 5 occurs here*]

This example highlights two crucial phenomena: firstly, the region pointed

at is not the region referred to. The indicated location in gesture space has to be mapped onto a spatial area of the described situation. Such mapping are accounted for by $v(\vec{p})$. Secondly, the gesture is produced *after* speech. This means that in order to identify the gesture's verbal attachment site something beyond mere temporal synchrony (the second basic question, see Section 3) has to be exploited. To this end, phonological information on top of the basic integration scheme in (6) are used (Alahverdzhieva & Lascarides 2011; Lücking 2013). Within a temporal window (basic integration scheme), a pointing gesture is affiliated to a stressed expression. This relaxation captures also post-speech gestures as in example (8)/Figure 5. Within HPSG, such constraints can be formulated within an interface to metrical trees from the phonological model of Klein (2000) or phonetic information packing from Engdahl & Vallduví (1996) – Tseng (2019), Chapter 22 of this volume. In order to deal with gestures that are affiliated with expressions that are larger than single words, Alahverdzhieva & Lascarides (2011) also develop a phrase or sentence level integration scheme, where the stressed element has to be a semantic head (in the study of Mehler & Lücking (2012) 18.8% of the gestures had a phrasal affiliate). On this account, semantic integration of gesture location and verbal meaning (the third basic question from Section 3) is brought about using underspecification mechanism of *Multiple Recursion Semantics* (MRS) (Copestake et al. 2005).

(Pointing) gestures that occur in company to a demonstrative expression in speech have received special attention, since they manifest the transition point between the symbolic realm of language and the deictic realm of perception (Levinson 2008; Fricke 2012). Building on this phenomenon, Ebert (2014) argue within in a dynamic semantic framework that co-demonstrative gestures contribute to at-issue content while other gestures contribute to non-at-issue content. Within a presuppositional framework, this picture has been refined due to gestures that are entailed in local contexts in terms of "cosuppositions" Schlenker (2018).

A dialogue-oriented focus is taken in Lücking (2018): here pointing gesture play a role in formulating processing instructions that guide the addressee in understanding demonstrative noun phrases (see also Lücking, Ginzburg & Cooper (2019), Chapter 30 of this volume).

Semantic aspects of iconic features of pointing as discourse referent indication in American Sign Language have been investigated and formally modelled by Schlenker et al. (2013); Schlenker (2014).

4.4 Iconic Gestures

There is nearly no semantic work on how and which meanings can be assigned to iconic gestures (first basic question from Section 3). Two exceptions being the approaches of Rieser (2010) and Lücking (2013; 2016). Rieser (2010) tries to extract a "depiction typology" out of a speech-and-gesture corpus where formal gesture features are correlated with topological clusters consisting of geometrical constructs. These geometrical objects are used in order to provide a possibly underspecified semantic representation for iconic gestures which is integrated into word meaning (Hahn & Rieser 2010; Rieser 2011). Lücking (2013; 2016) draws on Goodman's notion of *exemplification* (Goodman 1976), iconic gestures are connected to semantic predicates in terms of a reversed denotation relation: the meaning of an iconic gesture is given in terms of the set of predicates which have the gesture event within their denotation. In a further step, common perceptual features are extracted and represented as part of a lexical extension of lexemes. This conception is motivated by psychophysic theories of the perception of biological events (Johansson 1973), draws on philosophical similarity conception beyond isomorphic mappings (Peacocke 1987),[5] and, using a somewhat related approach at least, has been proven to work in robotics (Sowa 2006). These perceptual features serve as the integration locus for iconic gestures, using standard unification techniques. With regard to conventionalized sign language, iconic aspects have also been captured in terms of presuppositions that map, for instance, height specifications of signed loci to social properties and relations like power (that is, a higher locus indicates more power) (Schlenker et al. 2013).

The richer formal, functional and representational features of iconic gestures as compared to deictic gestures (cf. Section 4.3) is accounted for in Alahverdzhieva & Lascarides (2010) by assigning a formal predicate to each "phonological" feature of a gesture representation (cf. Section 4.2). These formal gesture predicates are highly underspecified, using *Robust Multiple Recursion Semantics* (RMRS) (Copestake 2007). That is, they can be assigned various (so far postulated) predicates (which are assumed to be constrained by iconicity (Alahverdzhieva & Lascarides 2010)) with differing arity in the gesture resolution process. Similar to pointing gesture integration, phonetic constraints on word and sentence level apply and semantic integration is achieved using underspecification techniques of *Multiple Recursion Semantics* (MRS) (Copestake et al. 2005). The received integrated representation are then resolved using the SDRT tools of Lascarides & Stone (2009).

---

[5]That mere resemblance, usually associated with iconic signs, is too empty a notion to provide the basis for a signifying relation has been emphasised on various occasions (Burks 1949; Bierman 1962; Eco 1976; Goodman 1976; Sonesson 1998).

Iconic gestures in particular are involved in a short-term dynamic phenomenon: on repeated co-occurrence, iconic gestures and affiliated speech can fuse into a *multimodal ensemble* (Kendon 2004; Lücking et al. 2008; Mehler & Lücking 2012). The characteristic feature of such an ensemble is that their gestural part, their verbal part, or even both parts can be simplified without changing the meaning of the ensemble. Ensembles, thus, are the result of a process sign formation as studied, for instance, in experimental semiotics Galantucci & Garrod (2011). Such grammaticalisation process eventually might lead to conventional signs. However, most conventional, emblematic everyday gesture seem to be the result of circumventing a taboo: something you should not name is gesticulated (Posner 2002).

5 Outlook

What are (still) challenging issues with respect to grammar-gesture integration, in particular from a semantic point of view? Candidates include:

- gestalt phenomena: the trajectories described by a gesture are often incomplete and have to be completed by drawing on gestalt principles or everyday knowledge (Lücking 2016);

- negligible features: not all formal features of a gesture are meaning-carrying features in the context of utterance. For instance, in a dynamic gesture the handshape often (though not always) does not provide any semantic information. How to distinguish between significant and negligible gesture features?

- "semantic endurance": due to holds, gestural meanings can kept fixed for some time is keeps available for semantic attachment. This may call for a more sophisticated algebraic treatment of speech-gesture integration than offered by typed feature structures (Rieser 2015).

Finally, the empirical domain of "gesture" has to extended to other non-verbal signals such as laughter (Ginzburg et al. 2015), facial expressions or gaze (see Section 1 for a brief list of non-verbal signals), in isolation as well as in mutual combination. Thus, there is still some way to go in order to achieve a fuller understanding of natural language interaction and thereby natural languages.

Acknowledgements

References

Abner, Natasha, Kensy Cooperrider & Susan Goldin-Meadow. 2015. Gesture for linguists: A handy primer. *Language and Linguistics Compass* 9(11). 437–451. DOI:10.1111/lnc3.12168

Alahverdzhieva, Katya. 2013. *Alignment of speech and co-speech gesture in a constraint-based grammar*. School of Informatics, University of Edinburgh dissertation.

Alahverdzhieva, Katya, Dan Flickinger & Alex Lascarides. 2012. Multimodal grammar implementation. In Eric Fosler-Lussier, Ellen Riloff & Srinivas Bangalore (eds.), *Proceedings of the 2012 conference of the North American Chapter of the Association for Computational Linguistics: human language technologies* (NAACL-HLT 2012), 582–586. Montreal, Canada.

Alahverdzhieva, Katya & Alex Lascarides. 2010. Analysing language and coverbal gesture in constraint-based grammars. In Stefan Müller (ed.), *Proceedings of the 17th International Conference on Head-Driven Phrase Structure Grammar, Université Paris Diderot*, 5–25. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2010/, accessed 2018-2-25.

Alahverdzhieva, Katya & Alex Lascarides. 2011. An HPSG approach to synchronous speech and deixis. In Stefan Müller (ed.), *Proceedings of the 18th International Conference on Head-Driven Phrase Structure Grammar, University of Washington*, 6–24. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2011/, accessed 2018-2-25.

Argyle, Michael. 1975. *Bodily communication*. New York, NY: Methuen & Co.

Argyle, Michael & Mark Cook. 1976. *Gaze and mutual gaze*. Cambridge, UK: Cambridge University Press.

Bangerter, Adrian & Daniel M. Oppenheimer. 2006. Accuracy in detecting referents of pointing gestures unaccompanied by language. *Gesture* 6(1). 85–102.

Bavelas, Janet B., Nicole Chovil, Linda Coates & Lori Roe. 1995. Gestures specialized for dialogue. *Personality and Social Psychology Bulletin* 21(4). 394–405.

Bavelas, Janet B., Nicole Chovil, Douglas A. Lawrie & Allan Wade. 1992. Interactive gestures. *Discourse Processes* 15(4). 469–489. DOI:10.1080/01638539209544823

Bavelas, Janet B., Jennifer Gerwing, Chantelle Sutton & Danielle Prevost. 2008. Gesturing on the telephone: Independent effects of dialogue and visibility. *Journal of Memory and Language* 58(2). 495–520. DOI:10.1016/j.jml.2007.02.004

Bender, Emily M. & Guy Emerson. 2019. Computational linguistics and grammar engineering. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 601–642. Berlin: Language Science Press. DOI:??

Bierman, Arthur K. 1962. That there are no iconic signs. *Philosophy and Phenomenological Research* 23(2). 243–249.

Birdwhistell, Ray L. 1970. *Kinesics and context: Essays on body motion communication.* Philadelphia: University of Pennsylvania Press.

Bruneau, Thomas J. 1980. Chronemics and the verbal-nonverbal interface. In Mary Ritchie Key (ed.), *The relationship of verbal and nonverbal communication*, 101–118. The Hague, The Netherlands: Mouton.

Bruner, Jerome S. 1998. Routes to reference. *Pragmatics & Cognition* 6(1/2). 209–227. Special Issue: The Concept of Reference in the Cognitive Sciences.

Bühler, Karl. 1934. *Sprachtheorie. Die Darstellungsfunktion der Sprache.* Stuttgart: Gustav Fischer Verlag. Re-edition Stuttgart: UTB, Lucius & Lucius, 1999.

Burks, Arthur W. 1949. Icon, index, and symbol. *Philosophy and Phenomenological Research* 9(4). 673–689.

Butterworth, George & Shoji Itakura. 2000. How the eyes, head and hand serve definite reference. *British Journal of Developmental Psychology* 18(1). 25–50.

Clark, Herbert H. 1996. *Using language.* Cambridge: Cambridge University Press.

Cohen, Philip R., Michael Johnston, David McGee, Sharon Oviatt, Jay Pittman, Ira Smith, Liang Chen & Josh Clow. 1997. QuickSet: Multimodal interaction for distributed applications. In *Proceedings of the fifth ACM international conference on multimedia* (MULTIMEDIA '97), 31–40. Seattle, Washington, USA. DOI:10.1145/266180.266328

Cooper, Robin & Jonathan Ginzburg. 2015. Type theory with records for natural language semantics. In Shalom Lappin & Chris Fox (eds.), *The handbook of contemporary semantic theory*, 2nd edn., chap. 12, 375–407. Oxford, UK: Wiley-Blackwell.

Cooperrider, Kensy. 2017. Foreground gesture, background gesture. *Gesture* 16(2). 176–202. DOI:10.1075/gest.16.2.02coo

Cooperrider, Kensy & Rafael Núñez. 2012. Nose-pointing: Notes on a facial gesture of Papua New Guinea. *Gesture* 12(2). 103–129. DOI:doi:10.1075/gest.12.2.01coo

Copestake, Ann. 2007. Semantic composition with (Robust) Minimal Recursion Semantics. In *Proceedings of the workshop on deep linguistic processing* (DeepLP'07), 73–80. Prague, Czech Republic.

Copestake, Ann, Dan Flickinger, Carl Pollard & Ivan A. Sag. 2005. Minimal Recursion Semantics: An introduction. *Research on Language and Computation* 3(4). 281–332.

Domaneschi, Filippo, Marcello Passarelli & Carlo Chiorri. 2017. Facial expressions and speech acts: Experimental evidences on the role of the upper face as an illocutionary force indicating device in language comprehension. *Cognitive Processing* 18(3). 285–306. DOI:10.1007/s10339-017-0809-6

Ebert, Cornelia. 2014. *The non-at-issue contributions of gestures.* Workshop on Demonstration and Demonstratives, April 11–12 2014, Stuttgart.

Eco, Umberto. 1976. *A theory of semiotics.* Bloomington: Indiana University Press.

Ekman, Paul & Wallace V. Friesen. 1969. The repertoire of nonverbal behavior: Categories, origins, usage, and coding. *Semiotica* 1(1). 49–98.

Ekman, Paul & Wallace V. Friesen. 1978. *Facial action coding system: A technique for the measurement of facial movement.* Palo Alto, CA: Consulting Psychologists Press.

Enfield, Nick J. 2001. Lip-pointing: A discussion of form and function with reference to data from Laos. *Gesture* 1(2). 185–212.

Engdahl, Elisabet & Enric Vallduví. 1996. Information packaging in HPSG. In Elisabet Engdahl & Enric Vallduví (eds.), *Edinburgh working papers in cognitive science* (Studies in HPSG 12), 1–31. Edinburgh: University of Edinburgh.

Fernando, Tim. 2011. Constructing situations and time. *Journal of Philosophical Logic* 40(3). 371–396. DOI:10.1007/s10992-010-9155-1

Fricke, Ellen. 2007. *Origo, Geste und Raum* (Linguistik – Impulse & Tendenzen 24). Berlin: De Gruyter.

Fricke, Ellen. 2012. *Grammatik multimodal. Wie Wörter und Gesten zusammenwirken* (Linguistik – Impulse und Tendenzen 40). Berlin: De Gruyter.

Galantucci, Bruno & Simon Garrod. 2011. Experimental semiotics: A review. *Frontiers in Human Neuroscience* 5(11). DOI:10.3389/fnhum.2011.00011

Gerwing, Jennifer & Janet B. Bavelas. 2004. Linguistic influences on gesture's form. *Gesture* 4(2). 157–195.

Gibbon, Dafydd, Ulrike Gut, Benjamin Hell, Karin Looks, Alexandra Thies & Thorsten Trippel. 2003. A computational model of arm gestures in conver-

sation. In ISCA Archive (ed.), *Proceedings of the 8th European conference on speech communication and technology* (EUROSPEECH 2003), 813–816. Geneva, Switzerland. http://www.isca-speech.org/archive/eurospeech_2003, accessed 2018-9-10.

Ginzburg, Jonathan, Ellen Breitholtz, Robin Cooper, Julian Hough & Ye Tian. 2015. Understanding laughter. In Thomas Brochhagen, Floris Roelofsen & Nadine Theiler (eds.), *Proceedings of the 20th Amsterdam Colloquium*, 137–146. Amsterdam, Netherlands. http://semanticsarchive.net/Archive/mVkOTk2N/AC2015-proceedings.pdf, accessed 2018-9-10.

Giorgolo, Gianluca. 2010. A formal semantics for iconic spatial gestures. In Maria Aloni, Harald Bastiaanse, Tikitu de Jager & Katrin Schulz (eds.), *Logic, language and meaning* (Lecture Notes in Computer Science 6042), 305–314. Berlin: Springer.

Giorgolo, Gianluca & Ash Asudeh. 2011. Multimodal communication in LFG: gestures and the correspondence architecture. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG 2011 conference*, 257–277. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/LFG/16/, accessed 2018-9-30.

Goodman, Nelson. 1976. *Languages of art. An approach to a theory of symbols*. 2nd edn. Idianapolis: Hackett Publishing Company, Inc.

Goodwin, Charles. 2003. Pointing as situated practice. In Sotaro Kita (ed.), *Pointing: Where language, culture, and cognition meet*, chap. 2, 217–241. Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc.

Hadjikhani, Nouchine, Rick Hoge, Josh Snyder & Beatrice de Gelder. 2008. Pointing with the eyes: The role of gaze in communicating danger. *Brain and Cognition* 68(1). 1–8.   DOI:http://dx.doi.org/10.1016/j.bandc.2008.01.008

Hahn, Florian & Hannes Rieser. 2010. Explaining speech gesture alignment in MM dialogue using gesture typology. In Paweł Łupowski & Matthew Purver (eds.), *Proceedings of the 14th workshop on the semantics and pragmatics of dialogue: Aspects of semantics and pragmatics of dialogue* (SemDial 2010), 99–111. Poznan, Poland.

Hall, Edward T. 1968. Proxemics. *Current Anthropology* 9(2-3). 83–108.

Jannedy, Stefanie & Norma Mendoza-Denton. 2005. Structuring information through gesture and intonation. *Interdisciplinary Studies on Information Structure (ISIS)* (3). 199–244. http://opus.kobv.de/ubp/volltexte/2006/877/, accessed 2018-9-30. URN: urn:nbn:de:kobv:517-opus-8774.

Johansson, Gunnar. 1973. Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics* 14(2). 201–211. DOI:10.3758/BF03212378

Johnston, Michael. 1998. Unification-based multimodal parsing. In *Proceedings of the 36th annual meeting of the Association for Computational Linguistics and 17th international conference on computational linguistics – Volume 1* (ACL '98), 624–630. Montreal, Quebec, Canada: Association for Computational Linguistics. DOI:10.3115/980845.980949

Johnston, Michael, Philip R. Cohen, David McGee, Sharon L. Oviatt, James A. Pittman & Ira Smith. 1997. Unification-based multimodal integration. In *Proceedings of the 35th annual meeting of the Association for Computational Linguistics and eighth conference of the european chapter of the association for computational linguistics*, 281–288. Madrid, Spain: Association for Computational Linguistics. DOI:10.3115/976909.979653

Kauffman, Lynn E. 1971. Tacesics, the study of touch: A model for proxemic analysis. *Semiotica* 4(2). 149–161. DOI:10.1515/semi.1971.4.2.149

Kelly, Spencer D., Corinne Kravitz & Michael Hopkins. 2004. Neural correlates of bimodal speech and gesture comprehension. *Brain and Language* 89(1). 253–260. DOI:10.1016/S0093-934X(03)00335-3

Kendon, Adam. 1967. Some functions of gaze-direction in social interaction. *Acta Psychologica* 26(1). 22–63. DOI:10.1016/0001-6918(67)90005-4

Kendon, Adam. 1972. Some relationships between body motion and speech. An analysis of an example. In Aron Wolfe Siegman & Benjamin Pope (eds.), *Studies in dyadic communication*, chap. 9, 177–210. Elmsford, NY: Pergamon Press.

Kendon, Adam. 1980. Gesticulation and speech: Two aspects of the process of utterance. In Mary Ritchie Key (ed.), *The relationship of verbal and nonverbal communication* (Contributions to the Sociology of Language 25), 207–227. The Hague: Mouton.

Kendon, Adam. 2004. *Gesture: Visible action as utterance.* Cambridge, MA: Cambridge University Press.

Kita, Sotaro & Aslı Özyürek. 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. DOI:10.1016/S0749-596X(02)00505-3

Klein, Ewan. 2000. Prosodic constituency in HPSG. In Ronnie Cann, Claire Grover & Philip Miller (eds.), *Grammatical interfaces in HPSG* (Studies in Constraint-Based Lexicalism), chap. 10, 169–200. Stanford, CA: CSLI Publications.

Klein, Wolfgang. 1978. Wo ist hier? Präliminarien zu einer Untersuchung der lokalen Deixis. *Linguistische Berichte* 58. 18–40.

Kopp, Stefan, Paul Tepper & Justine Cassell. 2004. Towards integrated microplanning of language and iconic gesture for multimodal output. In *Proceedings of the 6th international conference on multimodal interfaces* (ICMI '04), 97–104. State College, PA, USA: ACM. DOI:10.1145/1027933.1027952

Kranstedt, Alfred, Andy Lücking, Thies Pfeiffer, Hannes Rieser & Ipke Wachsmuth. 2006a. Deictic object reference in task-oriented dialogue. In Gert Rickheit & Ipke Wachsmuth (eds.), *Situated communication*, 155–207. Berlin: Mouton de Gruyter.

Kranstedt, Alfred, Andy Lücking, Thies Pfeiffer, Hannes Rieser & Ipke Wachsmuth. 2006b. Deixis: How to determine demonstrated objects using a pointing cone. In Sylvie Gibet, Nicolas Courty & Jean-Francois Kamp (eds.), *Gesture in human-computer interaction and simulation: 6th International Gesture Workshop, gw 2005, Berder Island, France, May 18-20, 2005, revised selected papers* (Lecture Notes in Computer Science 3881), 300–311. Berlin: Springer.

Krauss, Robert M., Yihsiu Chen & Rebecca F. Gottesmann. 2000. Lexical gestures and lexical access: A process model. In David McNeill (ed.), *Language and gesture*, chap. 13, 261–283. Cambridge, UK: Cambridge University Press.

Lascarides, Alex & Matthew Stone. 2009. A formal semantic analysis of gesture. *Journal of Semantics* 26(4). 393–449. DOI:10.1093/jos/ffp004

Lausberg, Hedda & Han Sloetjes. 2009. Coding gestural behavior with the NEUROGES-ELAN system. *Behavior Research Methods* 41(3). 841–849.

Levinson, Stephen C. 2008. Deixis. In *The handbook of pragmatics*, chap. 5, 97–121. Blackwell.

Loehr, Daniel. 2004. *Gesture and intonation*. Washington, D.C.: Georgetown University dissertation.

Loehr, Daniel. 2007. Apects of rhythm in gesture in speech. *Gesture* 7(2). 179–214.

Lücking, Andy. 2013. *Ikonische Gesten. Grundzüge einer linguistischen Theorie.* Berlin: De Gruyter. Zugl. Diss. Univ. Bielefeld (2011).

Lücking, Andy. 2016. Modeling co-verbal gesture perception in type theory with records. In Maria Ganzha, Leszek Maciaszek & Marcin Paprzycki (eds.), *Proceedings of the 2016 federated conference on computer science and information systems* (Annals of Computer Science and Information Systems 8), 383–392. IEEE. DOI:10.15439/2016F83

Lücking, Andy. 2018. Witness-loaded and witness-free demonstratives. In Marco Coniglio, Andrew Murphy, Eva Schlachter & Tonjes Veenstra (eds.), *Atypical*

demonstratives. Syntax, semantics and pragmatics (Linguistische Arbeiten 568). Berlin: De Gruyter.

Lücking, Andy, Kirsten Bergmann, Florian Hahn, Stefan Kopp & Hannes Rieser. 2010. The Bielefeld speech and gesture alignment corpus (SaGA). In *Multimodal corpora: Advances in capturing, coding and analyzing multimodality* (LREC 2010), 92–98. Malta.   DOI:10.13140/2.1.4216.1922

Lücking, Andy, Jonathan Ginzburg & Robin Cooper. 2019. Grammar in dialogue. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 643–676. Berlin: Language Science Press.   DOI:??

Lücking, Andy, Dennis Kurfürst, Désirée Walther, Marcel Mauri & Alexander Mehler. 2016. *FIGURE: Gesture annotation*. Tech. rep. TTLab-TR-2016-05. Goethe University Frankfurt: Text Technology Lab.

Lücking, Andy & Alexander Mehler. 2012. What's the scope of the Naming Game? Constraints on semantic categorization. In Thomas C. Scott-Phillips, Mónica Tamariz, Erica A. Cartmill & James R. Hurford (eds.), *Proceedings of the 9th international conference on the evolution of language* (Evolang IX), 196–203. Kyoto, Japan.

Lücking, Andy, Alexander Mehler & Peter Menke. 2008. Taking fingerprints of speech-and-gesture ensembles: Approaching empirical evidence of intrapersonal alignmnent in multimodal communication. In *Proceedings of the 12th workshop on the semantics and pragmatics of dialogue* (LonDial'08), 157–164. King's College London.

Lücking, Andy & Thies Pfeiffer. 2012. Framing multimodal technical communication. In Alexander Mehler & Laurent Romary (eds.). In collab. with Dafydd Gibbon, *Handbook of technical communication* (Handbooks of Applied Linguistics 8), chap. 18, 591–644. Berlin: De Gruyter Mouton.

Lücking, Andy, Thies Pfeiffer & Hannes Rieser. 2015. Pointing and reference reconsidered. *Journal of Pragmatics* 77. 56–79.   DOI:10.1016/j.pragma.2014.12.013

Lücking, Andy, Hannes Rieser & Jens Stegmann. 2004. Statistical support for the study of structures in multi-modal dialogue: Inter-rater agreement and synchronization. In Jonathan Ginzburg & Enric Vallduví (eds.), *Proceedings of the eighth workshop on the semantics and pragmatics of dialogue* (Catalog '04), 56–63. Barcelona.

Martell, Craig. 2002. FORM: an extensible, kinematically-based gesture annotation scheme. In Mark Maybury (ed.), *Proceedings of the third international conference on language resources and evaluation and the workshop on multimodal*

resources and multimodal systems evaluation (LREC 2002), 183–187. Las Palmas, Spain.

Masataka, Nobuo. 2003. From index-finger extension to index-finger pointing: Ontogenesis of pointing in preverbal infants. In Sotaro Kita (ed.), *Pointing: Where language, culture, and cognition meet*, chap. 4, 69–84. Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc.

Matthews, Danielle, Tanya Behne, Elena Lieven & Michael Tomasello. 2012. Origins of the human pointing gesture: A training study. *Developmental Science* 15(6). 817–829. DOI:10.1111/j.1467-7687.2012.01181.x

McClave, Evelyn. 1994. Gestural beats: The rhythm hypothesis. *Journal of Psycholinguistic Research* 23(1). 45–66.

McGinn, Colin. 1981. The mechanism of reference. *Synthese* 49(2). 157–186.

McNeill, David. 1985. So you think gestures are nonverbal? *Psychological Review* 92(3). 350–371.

McNeill, David. 1992. *Hand and mind – What gestures reveal about thought*. Chicago: Chicago University Press.

McNeill, David. 2005. *Gesture and thought*. Chicago: University of Chicago Press.

Mehler, Alexander & Andy Lücking. 2012. Pathways of alignment between gesture and speech: Assessing information transmission in multimodal ensembles. In Gianluca Giorgolo & Katya Alahverdzhieva (eds.), *Proceedings of the international workshop on formal and computational approaches to multimodal communication under the auspices of ESSLLI 2012, Opole, Poland, 6–10 August*.

Müller, Cornelia. 1998. *Redebegleitende Gesten. Kulturgeschichte – Theorie – Sprachvergleich* (Körper – Kultur – Kommunikation 1). Berlin: Berlin Verlag. Zugl. Diss. Frei Universität Berlin (1996).

Müller, Cornelia, Alan Cienki, Ellen Fricke, Silva Ladewig, David McNeill & Sedinha Tessendorf (eds.). 2013–2014. *Body – Language – Communication. An International Handbook on Multimodality in Human Interaction*. 2 vols. (Handbücher zur Sprach- und Kommunikationswissenschaft / Handbooks of Linguistics and Communication Science (HSK) 38/1 & 2). Berlin: De Gruyter Mouton.

Nobe, Shuichi. 2000. Where do *most* spontaneous representational gestures actually occur with respect to speech? In David McNeill (ed.), *Language and gesture*, 186–198. Cambridge, MA: Cambridge University Press.

Nöth, Winfried. 1990. *Handbook of semiotics*. Bloomington, Indianapolis: Indiana University Press.

Nunberg, Geoffrey. 1993. Indexicality and deixis. *Linguistics and Philosophy* 16(1). 1–43. DOI:10.1007/BF00984721

Draft of 1st April 2019, 11:14

Özyürek, Aslı. 2012. Gesture. In Roland Pfau, Markus Steinbach & Bencie Woll (eds.), *Sign language: An international handbook*, vol. 37 (Handbücher zur Sprach- und Kommunikationswissenschaft / Handbooks of Linguistics and Communication Science (HSK)), chap. 27, 626–646. Berlin: De Gruyter Mouton.

Peacocke, Christopher. 1987. Depiction. *The Philosophical Review* 96(3). 383–410.

Posner, Roland. 2002. Alltagsgesten als Ergebnis von Ritualisierung. In Jan C. Joerden (ed.), *Ritualisierte Tabuverletzung, Lachkultur und das Karnevaleske*, vol. 6 (Studien zur Ethik in Ostmitteleuropa), 395–421. Frankfurt am Main: Peter Lang.

Posner, Roland, Klaus Robering & Thomas A. Sebeok (eds.). 1997–2004. *Semiotik. Ein Handbuch zu den zeichentheoretischen Grundlagen von Natur und Kultur / Semiotics. A Handbook on the Sign-Theoretic Foundations of Nature and Culture.* 4 vols. (Handbücher zur Sprach- und Kommunikationswissenschaft / Handbooks of Linguistics and Communication Science (HSK) 13/1–4). Berlin: De Gruyter.

Poyatos, Fernando. 1975. Cross-cultural study of paralinguistic "alternants" in face-to-face interaction. In Adam Kendon, Harris M. Richards & Mary Ritchie Key (eds.), *Organization of behavior in face-to-face interaction*, 285–314. The Hague: Mouton.

Quine, Willard Van Orman. 1950. Identity, ostension, and hypostasis. *The Journal of Philosophy* 47(22). 621–633.

Rieser, Hannes. 2004. Pointing in dialogue. In *Proceedings of the eighth workshop on the semantics and pragmatics of dialogue* (Catalog '04), 93–100. Barcelona.

Rieser, Hannes. 2008. Aligned iconic gesture in different strata of mm route-description. In *LonDial 2008: The 12th workshop on the semantics and pragmatics of dialogue* (SemDial), 167–174. King's College London.

Rieser, Hannes. 2010. On factoring out a gesture typology from the Bielefeld Speech-Gesture-Alignment Corpus. In Stefan Kopp & Ipke Wachsmuth (eds.), *Proceedings of GW 2009: Gesture in embodied communication and human-computer interaction*, 47–60. Berlin & Heidelberg: Springer.

Rieser, Hannes. 2011. How to disagree on a church window's shape using gesture. In Klaus Hölker & Carla Marello (eds.), *Dimensionen der Analyse von Texten und Diskursen. Festschrift für János Sándor Petöfi*, 231–246. Berlin: LIT Verlag.

Rieser, Hannes. 2015. When hands talk to mouth. Gesture and speech as autonomous communicating processes. In Christine Howes & Staffan Larsson (eds.), *Proceedings of the 19th workshop on the semantics and pragmatics of dialogue* (SemDial 2015: goDIAL), 122–130. Gothenburg, Sweden.

Rieser, Hannes & Massimo Poesio. 2009. Interactive gestures in dialogue: A PTT model. In *Proceedings of the 10th annual meeting of the special interest group in discourse and dialogue* (SIGDIAL 2009), 87–96. Queen Mary University of London.

de Ruiter, Jan Peter. 2000. The production of gesture and speech. In David Mc-Neill (ed.), *Language and gesture*, chap. 14, 284–311. Cambridge, UK: Cambridge University Press.

de Ruiter, Jan Peter. 2004. On the primacy of language in multimodal communication. In Jean-Claude Martin, Elisabeth Den, Peter Kühnlein, Lou Boves, Patrizia Paggio & Roberta Catizone (eds.), *Proceedings of the fourth international conference on language resources and evaluation and the workshop on multimodal corpora: Models of human behaviour for the specification and evaluation of multimodal input and output interfaces* (LREC 2004), 38–41. Lisbon, Portugal.

de Ruiter, Jan Peter. 2007. Postcards from the mind: The relationship between speech, imagistic gesture, and thought. *Gesture* 7(1). 21–38.

Schlenker, Philippe. 2014. Iconic features. *Natural Language Semantics* 22(4). 299–356. DOI:10.1007/s11050-014-9106-4

Schlenker, Philippe. 2018. Gesture projection and cosuppositions. *Linguistics and Philosophy* 41(3). 295–365. DOI:10.1007/s10988-017-9225-8

Schlenker, Philippe, Jonathan Lamberton & Mirko Santoro. 2013. Iconic variables. *Linguistics and Philosophy* 36(2). 91–149. DOI:10.1007/s10988-013-9129-1

Slama-Cazacu, Tatiana. 1976. Nonverbal components in message sequence: "Mixed Syntax". In William C. McCormick & Stephan A. Wurm (eds.), *Language and man. Anthropological issues* (World Anthropology), 217–227. The Hague & Paris: Mouton.

van der Sluis, Ielka & Emiel Krahmer. 2007. Generating multimodal references. *Discourse Processes* 44(3). 145–174. Special Issue: Dialogue Modelling: Computational and Empirical Approaches. DOI:10.1080/01638530701600755

Sonesson, Göran. 1998. That there are many kinds of iconic signs. *Visio* 1(1). 33–54.

Sowa, Timo. 2006. *Understanding coverbal iconic gestures in shape descriptions*. Berlin: Akademische Verlagsgesellschaft. Zugl. Diss. Univ. Bielefeld.

Steels, Luc. 1995. A self-organizing spatial vocabulary. *Artificial Life* 2(3). 319–332.

Steinbach, Markus & Anke Holler. 2019. Sign languages. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 677–678. Berlin: Language Science Press. DOI:??

Streeck, Jürgen. 2008. Depicting by gesture. *Gesture* 8(3). 285–301. DOI:10.1075/gest.8.3.02str

Tseng, Jesse. 2019. Phonology. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 481–484. Berlin: Language Science Press. DOI:??

Tuite, Kevin. 1993. The production of gesture. *Semiotica* 93(1/2). 83–105.

Wagner, Petra, Zofia Malisz & Stefan Kopp. 2014. Gesture and speech in interaction: An overview. *Speech Communication* 57. 209–232. DOI:10.1016/j.specom.2013.09.008

Wilkins, David. 2003. Why pointing with the index finger is not a universal (in sociocultural and semiotic terms). In Sotaro Kita (ed.), *Pointing: Where language, culture, and cognition meet*, chap. 8, 171–215. Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc.

Wiltshire, Anne. 2007. Synchrony as the underlying structure of gesture: The relationship between speech sound and body movement at the micro level. In Robyn Loughnane, Cara Penry Williams & Jana Verhoeven (eds.), *In between wor(l)ds: Transformation and translation*, vol. 6 (Postgraduate Research Papers on Language and Literature), 235–253. Melbourne: School of Languages & Linguistics, The University of Melbourne.

Wu, Ying Choon & Seanna Coulson. 2005. Meaningful gestures: electrophysiological indices of iconic gesture comprehension. *Psychophysiology* 42(6). 654–667.

Part V

The broader picture