2.	$\begin{bmatrix} \textsc{spkr} & = & B \\ \textsc{addr} & = & A \\ \textsc{Moves} & = & \left\langle \textsc{Ask}(\text{B,A,}\textsc{q}(2)),\ \textsc{Assert}(\text{A,B,}\textsc{p}(1)) \right\rangle \\ \textsc{qud} & = & \left\langle \textsc{q}(2) \right\rangle \\ \textsc{facts} & = & \mathit{cg0} \cup \left\{ \textsc{p}(1) \right\} \end{bmatrix}$	“Which university?” Accept + Ask QUD-incrementation
3.	$\begin{bmatrix} \textsc{spkr} & = & A \\ \textsc{addr} & = & B \\ \textsc{Moves} & = & \left\langle \begin{array}{l} \textsc{Assert}(\text{A,B,}\textsc{p}(3)), \\ \textsc{Ask}(\text{B,A,}\textsc{q}(2)),\ \textsc{Assert}(\text{A,B,}\textsc{p}(1)) \end{array} \right\rangle \\ \textsc{qbg} & = & \mathit{About(p(3),q(2))} \\ \textsc{qud} & = & \left\langle \textsc{p}?(3),\ \textsc{q}(2) \right\rangle \\ \textsc{facts} & = & \mathit{cg0} \cup \left\{ \textsc{p}(1) \right\} \end{bmatrix}$	“Cambridge.” QSPEC (via *About* relation) + Assert QUD-incrementation
4.	$\begin{bmatrix} \textsc{spkr} & = & B \\ \textsc{addr} & = & A \\ \textsc{Moves} & = & \left\langle \begin{array}{l} \textsc{Accept}(\text{B,A,}\textsc{p}(3)),\ \textsc{Assert}(\text{A,B,}\textsc{p}(3)), \\ \textsc{Ask}(\text{B,A,}\textsc{q}(2)),\ \textsc{Assert}(\text{A,B,}\textsc{p}(1)) \end{array} \right\rangle \\ \textsc{qud} & = & \left\langle \quad \right\rangle \\ \textsc{facts} & = & \mathit{cg0} \cup \left\{ \textsc{p}(3),\ \textsc{p}(1) \right\} \end{bmatrix}$	“Cambridge, um.” Accept + Fact update/QUD-downdate
5.	$\begin{bmatrix} \textsc{spkr} & = & B \\ \textsc{addr} & = & A \\ \textsc{Moves} & = & \left\langle \begin{array}{l} \textsc{Ask}(\text{B,A,}\textsc{q}(5)), \\ \textsc{Accept}(\text{B,A,}\textsc{p}(3)),\ \textsc{Assert}(\text{A,B,}\textsc{p}(3)), \\ \textsc{Ask}(\text{B,A,}\textsc{q}(2)),\ \textsc{Assert}(\text{A,B,}\textsc{p}(1)) \end{array} \right\rangle \\ \textsc{qud} & = & \left\langle \textsc{q}(5) \right\rangle \\ \textsc{facts} & = & \mathit{cg0} \cup \left\{ \textsc{p}(3),\ \textsc{p}(1) \right\} \end{bmatrix}$	“what did you read?” Free Speech + Ask QUD-incrementation
6.	$\begin{bmatrix} \textsc{spkr} & = & A \\ \textsc{addr} & = & B \\ \textsc{Moves} & = & \left\langle \begin{array}{l} \textsc{Assert}(\text{A,B,}\textsc{p}(6)),\ \textsc{Ask}(\text{B,A,}\textsc{q}(5)), \\ \textsc{Accept}(\text{B,A,}\textsc{p}(3)),\ \textsc{Assert}(\text{A,B,}\textsc{p}(3)), \\ \textsc{Ask}(\text{B,A,}\textsc{q}(2)),\ \textsc{Assert}(\text{A,B,}\textsc{p}(1)) \end{array} \right\rangle \\ \textsc{qbg} & = & \mathit{About(p(6),q(5))} \\ \textsc{qud} & = & \left\langle \textsc{p}?(6),\ \textsc{q}(5) \right\rangle \\ \textsc{facts} & = & \mathit{cg0} \cup \left\{ \textsc{p}(3),\ \textsc{p}(1) \right\} \end{bmatrix}$	“History and English.” QSPEC (via *About* relation) + Assert QUD-incrementation

7.	$\left[\begin{array}{lll}\text{SPKR} & = & B \\ \text{ADDR} & = & A \\ \text{MOVES} & = & \left\langle\begin{array}{l}\text{ACCEPT(B,A,P(6))}, \\ \text{ASSERT(A,B,P(6)), ASK(B,A,Q(5))}, \\ \text{ACCEPT(B,A,P(3)), ASSERT(A,B,P(3))}, \\ \text{ASK(B,A,Q(2)), ASSERT(A,B,P(1))}\end{array}\right\rangle \\ \text{QUD} & = & \langle\ \rangle \\ \text{FACTS} & = & cg0 \cup \{\text{P(6), P(3), P(1)}\}\end{array}\right]$	"History and English." Accept + Fact update/QUD-downdate

Note that the dialogical exchange leads to an increase of the common ground of the interlocutors A and B: after chatting, the common ground contains the propositions *that A has been at university* (p(1)), *that A has been at Cambridge University* (p(3)) and *that A read History and English* (p(6)).

On these grounds, a lexical entry for "hello" can be spelled out. "Hello" realises a greeting move (which is its content) and must be used discourse-initially (the MOVES list and the QUD set have to be empty):

(24) $\left[\begin{array}{l}\text{PHON} : \langle \text{ HELLO } \rangle \\ \text{CAT} : \left[\text{HEAD}{=}\textit{INTERJECTION} \ : \ \textit{syncat}\right] \\ \text{DGB-PARAMS} : \left[\begin{array}{ll}\text{SPKR} & : \ \textit{Ind} \\ \text{ADDR} & : \ \textit{Ind} \\ \text{MOVES}{=}\langle\ \rangle & : \ \textit{list(IllocProp)} \\ \text{QUD}{=}\{\} & : \ \textit{poset(Question)}\end{array}\right] \\ \text{CONT}{=}\text{GREET(SPKR,ADDR)}: \textit{ILLOCPROP}\end{array}\right]$

Discourse-dynamically, "hello" puts a greeting move onto the MOVES list of the dialogue gameboard, thereby initiates an interaction and invites for a countergreeting (the requirement for countergreeting is exactly that a greeting move is the element of the otherwise empty list of dialogue moves) – giving rise to an *adjacency pair* as part of the local management system for dialogues investigated in conversational analysis (Schegloff & Sacks 1973).

The discourse particle "yes" can be used to answer a polar yes/no question. In this use, "yes" has a propositional content p that asserts the propositional content of the polar question p?, which has to be the maximal element in QUD (Ginzburg 2012: Chapter 2, 231 *et seq.*). That is, "yes" affirmatively resolves a given polar question. Polar questions, in turn, are 0-ary propositional abstracts (Ginzburg 2012: 231), that is, the polar question p? corresponding to a proposition p is a function mapping an empty record to p: $\lambda r : [].p$. Thus, applying p? to an empty record [] returns p, which is exactly what "yes" does. The affirmative particle (used to answer a yes/no question) is a propositional lexeme which applies a polar question which is maximal in QUD to an empty record (cf. Ginzburg 2012: 232):

(25)
$$\begin{bmatrix} \text{PHON} : \langle \text{ YES } \rangle \\ \text{CAT} : \begin{bmatrix} \text{HEAD}=\textit{PARTCL} \ : \ \textit{syncat} \end{bmatrix} \\ \text{DGB-PARAMS} : \begin{bmatrix} \text{QUD}=\begin{bmatrix} \text{MAX} & : \textit{PolQuestion} \\ \text{REST} & : \textit{set(Question)} \end{bmatrix} : \textit{poset(Question)} \end{bmatrix} \\ \text{CONT}=\text{DGB-PARAMS.QUD.MAX}(\ [\ ]\ ) : \textit{Prop} \end{bmatrix}$$

Due to its involvedness in DGB-PARAMS.QUD, "yes" directly interacts with accept and downdating, as described above. For more on this, see Ginzburg (2012).

5 Outlook

Given a basic framework for formulating and analysing content in dialogue context, there are various directions to explore, including the following ones.

- One of the main challenges of dialogue semantics is the integration of *non-verbal communication* means, like gaze, gestures, body posture, timing and non-language vocal sounds (e.g. laughter; Ginzburg et al. 2015; Tian et al. 2016). Since non-verbal communication means are informative, not only does a (dialogue) semantic representation have to be developed, but also the rules of their interaction with speech have to be formulated.

- Strictly speaking, dialogue is the interaction between *two* interlocutors. How can one scale up to *multilogue*, where the number of participants is at least three (Ginzburg & Fernández 2005)? Given the increased number of participants, problems that emerge include *grounding by proxy*, where a representative represents the dialogue gameboard of a group (Eshghi & Healey 2016) and of course *turn taking*.

- People do not process natural language input sentence-wise. Rather, processing begins with the initial sound and proceeds word for word or even on smaller units like affixes and phonemes – that is, processing is incremental (e.g. Sedivy et al. 1999; see also Wasow 2019, Chapter 27 of this volume). This is a key ingredient in the efficient (relatively gap-free and interruption-less) managing of turn taking. One direction of dialogue theories therefore is to bring psycholinguistics and formal semantics closer together by devising incremental grammar and dialogue gameboard models (Hough et al. 2015; Demberg et al. 2013; Poesio & Rieser 2011).

Finally, we want to mention two other dialogue-theoretic frameworks that have been worked out to a substantial degree, namely PTT (Traum 1994; Poesio 1995; Poesio & Traum 1997; Poesio & Rieser 2010), and *Segmented Discourse Representation Theory* (SDRT) (Asher 1993; Asher & Lascarides 2003; 2013; Hunter & Asher 2015). The phenomena and outlook directions discussed in this chapter apply to all theories of dialogue semantics, of course.

Appendix: An HPSG_{TTR} fragment

The appendix provides a fragment of HPSG_{TTR}. The grammar framework used is oriented at a *Head-driven Phrase Structure Grammar* variant (Sag et al. 2003), namely its TTR implementation (Cooper 2008). We use HPSG because its architecture satisfies the property of *incremental correspondence* (Johnson & Lappin 1999) – utterance representations encode phonological, syntactic, semantic and contextual information *fractally.* This is crucial *inter alia* for any treatment of clarification interaction (cf. Section 3.1). We use HPSG_{TTR} because the type-theoretical version allows us to directly incorporate semantic objects (cf. Section 3.2).

TTR has a counterpart to unification, namely the *merge* construction.

(26) a. If R_1 and R_2 are record types, then $R_1 \wedge_{merge} R_2$ is a record type and is called the *merge* of R_1 and R_2.

b. Since merge types are complicated to define (but see Cooper 2012), we follow the strategy of Cooper (2017) and illustrate the working of merges by means of some examples:

(i) $$\begin{bmatrix} \text{A} : T \\ \text{B} : R \end{bmatrix} \wedge_{MERGE} \begin{bmatrix} \text{C} : S \end{bmatrix} = \begin{bmatrix} \text{A} : T \\ \text{B} : R \\ \text{C} : S \end{bmatrix}$$

(ii) $$\begin{bmatrix} \text{A} : T \end{bmatrix} \wedge_{MERGE} \begin{bmatrix} \text{A} : R \end{bmatrix} = \begin{bmatrix} \text{A} : T \wedge_{merge} R \end{bmatrix}$$

Structure sharing is indicated by a "tag type" notation. Tag types are defined in terms of manifest fields.[8] The notational convention is exemplified in (27) by means of head-specifier agreement, where the tag type from (27a) abbreviates the structure in (27b):

[8] *NB:* technically, tag types apply singleton types to record types, instead of to objects, thereby making use of a revision of the notion of singleton types introduced by Cooper (2013: 4, footnote 3).

(27) a. $\left[\text{CAT} : \begin{bmatrix} \text{HEAD} : \left[\text{AGR}_{\boxed{1}} : Agr\right] \\ \text{SPR} : \left\langle\left[\text{CAT} : \left[\text{HEAD} : \left[\text{AGR=}\boxed{1} : Agr\right]\right]\right]\right\rangle \end{bmatrix}\right]$

b. $\left[\text{CAT} : \begin{bmatrix} \text{HEAD} : \left[\text{AGR} : Agr\right] \\ \text{SPR} : \left\langle\left[\text{CAT} : \left[\text{HEAD} : \left[\text{AGR=/CAT.HEAD.AGR} : Agr\right]\right]\right]\right\rangle \end{bmatrix}\right]$

The tag type notation alludes to the box notation common in HPSG work.

Agr is defined as usual:

(28) $Agr \coloneqq \begin{bmatrix} \text{NUM} : Num \\ \text{PERS} : Per \\ \text{GEN} : Gen \end{bmatrix}$

A basic *sign* is a pairing of phonetic, syntactic and semantic information and follows the geometry in (29):

(29) $sign \coloneqq \begin{bmatrix} \text{PHON} : Phoneme \\ \text{CAT} : SynCat \\ \text{DGB-PARAMS} : RecType \\ \text{CONT} : SemObj \end{bmatrix}$

Signs employ DGB-PARAMS, which host referential meanings that are witnessed among interlocutors. Quantificational abstraction is achieved by coercing parts of DGB-PARAMS to Q-PARAMS:

(30) If DGB-PARAMS : R_2 and for two record types R_0 and R_1 lacking any mutual dependencies[9] $R_2 = R_0 \wedge_{merge} R_1$, then R_0 can be moved to Q-PARAMS, resulting in the following structure:

$\begin{bmatrix} \text{DGB-PARAMS} : R_1 \\ \text{CONT} = \left[\text{Q-PARAMS} : R_0\right] \end{bmatrix}$

A word is a sign with constituent type (CXTYPE) *word*. Using the merge operation, the word extension on signs can represented compactly as in (31a), which expands to the structure given in (31b):

(31) a. $word \coloneqq sign \wedge_{merge} \left[\text{CXTYPE} : word\right] : RecType$

b. $\begin{bmatrix} \text{CXTYPE} : word \\ \text{PHON} : Phoneme \\ \text{CAT} : SynCat \\ \text{DGB-PARAMS} : RecType \\ \text{CONT} : SemObj \end{bmatrix}$

[9]None of the labels occurring in R_0 occur in R_1 and vice versa.

Words – that is, cxtype *word* – are usually the result of lexical rules, whose input are lexemes. Lexemes differ from words in their constituent type:

(32) $lexeme := sign \wedge_{merge} \left[\textsc{cxtype} : lexeme\right] : RecType$

A phrasal sign can be seen as a word with daughters:

(33) a. $phrase := sign \wedge_{merge} \left[\begin{array}{lll} \textsc{cxtype} & : & phrase \\ \textsc{dtrs} & : & \left[\textsc{nhd-dtrs} : List(Sign)\right] \end{array}\right] : RecType$

b. $$\left[\begin{array}{lll} \textsc{cxtype} & : & phrase \\ \textsc{phon} & : & List(Phoneme) \\ \textsc{cat} & : & SynCat \\ \textsc{dgb-params} & : & RecType \\ \textsc{cont} & : & SemObj \\ \textsc{dtrs} & : & \left[\textsc{nhd-dtrs} : List(Sign)\right] \end{array}\right]$$

A headed phrase is a phrase with a prominent daughter, i.e. the head daughter:

(34) a. $hd\text{-}phrase := phrase \wedge_{merge} \left[\textsc{dtrs} : \left[\textsc{hd-dtr} : Sign\right]\right] : RecType$

b. $$\left[\begin{array}{lll} \textsc{cxtype} & : & phrase \\ \textsc{phon} & : & List(Phoneme) \\ \textsc{cat} & : & SynCat \\ \textsc{dgb-params} & : & RecType \\ \textsc{cont} & : & SemObj \\ \textsc{dtrs} & : & \left[\begin{array}{lll} \textsc{hd-dtr} & : & Sign \\ \textsc{nhd-dtrs} & : & List(Sign) \end{array}\right] \end{array}\right]$$

The head daughter is special since it (as a default, at least) determines the syntactic properties of the mother construction. This aspect of headedness is captured in terms of the *head-feature principle* (HFP), which can be implemented by means of tag types as follows:

(35) $$\text{HFP} := \left[\begin{array}{lll} \textsc{cxtype} & : & phrase \\ \textsc{cat} & : & \left[\textsc{head}_{\boxed{2}} : PoS\right] \\ \textsc{dtrs} & : & \left[\textsc{hd-dtr} : \left[\textsc{cat} : \left[\textsc{head}{=}\boxed{2} : PoS\right]\right]\right] \end{array}\right]$$

The fact that the daughters' locutions combine to the mother's utterance is captured in terms of a "phon principle" (we use a slash notation in order to indicate paths starting at the outermost level of a feature structure):

(36) $$\text{PHON} := \left[\begin{array}{lll} \textsc{cxtype} & : & phrase \\ \textsc{phon} & : & List(/dtrs.hd.dtr/phon,\ /dtrs.nhd.dtrs/pos1.phon, \\ & & \ldots,\ /dtrs.nhd.dtrs/posn.phon) \end{array}\right]$$

Since semantic composition rests on predication rather than unification, there is no analog to the semantic compositionality principle of Sag et al. (2003) in our account. There is, however, something akin to semantic inheritance: we need to keep track of the contextual and quantificational paramaters contributed by the daughters of a phrase. This is achieved in terms of a *dgb-params principle* (*DGBPP*) in (37) which unifies the daughters' DGB-PARAMS into the mother's DGB-PARAMS (see Ginzburg 2012: 126 *et seq.* for a similar principle):

(37) DGBPP :=

$$\begin{bmatrix} \textsc{cxtype} & : & \textit{phrase} \\ \textsc{dgb-params} & : & \begin{bmatrix} /\textsc{dtrs.hd-dtr.dgb-params} \wedge_{\textit{MERGE}} /\textsc{dtrs.nhd-dtrs.pos1.dgb-params} \wedge_{\textit{MERGE}} \\ \ldots\wedge_{\textit{MERGE}} /\textsc{dtrs.nhd-dtrs.pos}n\textsc{.dgb-params} \end{bmatrix} \\ \textsc{dtrs} & : & \begin{bmatrix} \textsc{hd-dtr} & : & \begin{bmatrix} \textsc{q-params} : \textit{RecType} \end{bmatrix} \\ \textsc{nhd-dtrs} & : & \left\langle \textsc{pos1} : \begin{bmatrix} \textsc{q-params} : \textit{RecType} \end{bmatrix}, \ldots, \textit{posn} : \begin{bmatrix} \textsc{q-params} : \textit{RecType} \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix}$$

A headed phrase is well-formed when it is a headed phrase and it obeys the head feature principle, the phon principle and the dgb-params principle, which is expressed by extending *hd-phrase* by the following constraint:

(38) $\textit{hd-phrase} \coloneqq \textit{hd-phrase} \wedge_{\textit{merge}} \text{HFP} \wedge_{\textit{merge}} \text{PHON} \wedge_{\textit{merge}} \text{DGBPP}$

Using this set-up, lexical entries, lexical rules and syntactic constructions can be formulated straightforwardly.

References

Arnold, Doug. 2019. HPSG and Lexical Functional Grammar. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, i–ii. Berlin: Language Science Press. DOI:??

Asher, Nicholas. 1993. *Reference to abstract objects in discourse* (Studies in Linguistics and Philosophy 50). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Asher, Nicholas & Alex Lascarides. 2003. *Logics of conversation*. Cambridge: Cambridge University Press.

Asher, Nicholas & Alex Lascarides. 2013. Strategic conversation. *Semantics and Pragmatics* 6(2). 1–62. DOI:10.3765/sp.6.2

Austin, John L. 1950. Truth. In *Proceedings of the Aristotelian Society. Supplementary*, vol. xxiv, 111–128. Reprinted in John L. Austin: *Philosophical Papers*. 2. ed. Oxford: Clarendon Press, 1970.

Austin, John L. 1962. *How to do things with words*. 2nd edn. Cambridge, MA: Harvard University Press.

Barwise, Jon & John Perry. 1983. *Situations and attitudes* (The David Hume Series of Philosophy and Cognitive Science Reissues). Stanford: CSLI Publications.

Beißwenger, Michael, Maria Ermakova, Alexander Geyken, Lothar Lemnitzer & Angelika Storrer. 2012. DeRiK: A German reference corpus of computer-mediated communication. In *Proceedings of Digital Humanities 2012*. http://www.dh2012.uni-hamburg.de/conference/programme/abstracts/derik-a-german-reference-corpus-of-computer-mediated-communication/.

Blank, Malou, Anna Tornberg & Corina Knipper. 2018. New perspectives on the late neolithic of south-western Sweden. An interdisciplinary investigation of the gallery grave Falköping Stad 5. *Open Archaeology* 4(1). 1–35.

Bonami, Olivier. 2002. A syntax-semantics interface for tense and aspect in french. In Frank Van Eynde, Lars Hellan & Dorothee Beermann (eds.), *The proceedings of the 8th international conference on head-driven phrase structure grammar*, 31–50. Stanford: CSLI Publications. http://cslipublications.stanford.edu/HPSG/2/.

Borsley, Robert D. & Stefan Müller. 2019. HPSG and Minimalism. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, i–lvii. Berlin: Language Science Press. DOI:??

Carpenter, Bob. 1992. *The logic of typed feature structures*. Cambridge, MA: Cambridge University Press.

Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.

Clark, Herbert H. 1975. Bridging. In *Proceedings of the 1975 workshop on theoretical issues in natural language processing* (TINLAP '75), 169–174. Cambridge, Massachusetts: Association for Computational Linguistics. DOI:10.3115/980190.980237

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

Clark, Herbert H. & Catherine R. Marshall. 1981. Definite reference and mutual knowledge. In Aravind K. Joshi, Bonnie L. Webber & Ivan A. Sag (eds.), *Elements of discourse understanding*, 10–63. Cambridge: Cambridge University Press.

Clark, Herbert H., Robert Schreuder & Samuel Buttrick. 1983. Common ground and the understanding of demonstrative reference. *Journal of Verbal Learning and Verbal Behavior* 22(2). 245–258.

Cohen, Philip R. & Hector J. Levesque. 1990. Rational interaction as the basis for comunication. In Philip R. Cohen, Jerry R. Morgan & Martha E. Pollack (eds.), *Intentions in communication*, 221–256. Cambridge, MA: MIT Press.

Cooper, Robin. 2005a. Austinian truth, attitudes and type theory. *Research on Language and Computation* 3(2-3). 333–362. DOI:10.1007/s11168-006-0002-z

Cooper, Robin. 2005b. Records and record types in semantic theory. *Journal of Logic and Computation* 15(2). 99–112.

Cooper, Robin. 2008. Type theory with records and unification-based grammar. In Fritz Hamm & Stephan Kepser (eds.), *Logics for linguistic structures* (Trends in Linguistics: Studies and Monographs 201), 9–33. Berlin & New York: Mouton de Gruyter.

Cooper, Robin. 2012. Type theory and semantics in flux. In Ruth Kempson, Tim Fernando & Nicholas Asher (eds.), *Philosophy of linguistics* (Handbook of Philosophy of Science 6), 271–323. Oxford & Amsterdam: Elsevier.

Cooper, Robin. 2013. Clarification and generalized quantifiers. *Dialogue & Discourse* 4(1). 1–25. http://elanguage.net/journals/dad/article/view/1554, accessed 2015-1-13. DOI:10.5087/dad.2013.101

Cooper, Robin. 2014. Phrase structure rules as dialogue update rules. In *Procceedings of DialWatt – The 18th workshop on the semantics and pragmatics of dialogue* (SemDial 2014), 26–34. Edinburgh.

Cooper, Robin. 2017. Adapting Type Theory with Records for natural language semantics. In Stergios Chatzikyriakidis & Zhaohui Luo (eds.), *Modern perspectives in Type-Theoretical Semantics* (Studies in Linguistics and Philosophy 98), 71–94. Cham: Springer International Publishing AG. DOI:10.1007/978-3-319-50422-3_4

Cooper, Robin. 2019. *From perception to linguistic communication: Using a theory of types with records (TTR) to model linguistic action and content.* https://github.com/robincooper/ttl. MS Gothenburg University.

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.

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.

Costa, Francisco & António Branco. 2012. Backshift and tense decomposition. In Stefan Müller (ed.), *Proceedings of the 19th International Conference on Head-Driven Phrase Structure Grammar, Chungnam National University Daejeon*, 86–

106. CSLI Publications. http://cslipublications.stanford.edu/HPSG/2012/costa-branco.pdf.

Demberg, Vera, Frank Keller & Alexander Koller. 2013. Incremental, predictive parsing with psycholinguistically motivated Tree-Adjoining Grammar. *Computational Linguistics* 39(4). 1025–1066.

Devlin, Keith. 1991. *Logic and information*. Cambridge: Cambridge University Press.

Donnellan, Keith S. 1966. Reference and definite descriptions. *The Philosophical Review* 75(3). 281–304.

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.

Eshghi, Arash & Patrick G. T. Healey. 2016. Collective contexts in conversation: grounding by proxy. *Cognitive Science* 40(2). 299–324. DOI:10.1111/cogs.12225

Fernández, Raquel & Jonathan Ginzburg. 2002. Non-sentential utterances: A corpus study. *Traîtement Automatique de Languages* 43(2). 13–42.

Fernández, Raquel, Jonathan Ginzburg & Shalom Lappin. 2007. Classifying non-sentential utterances in dialogue: a machine learning approach. *Computational Linguistics* 33(3). 397–427. DOI:10.1162/coli.2007.33.3.397

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

Frege, Gottlob. 1892. Über Sinn und Bedeutung. In Mark Textor (ed.), *Funktion – Begriff – Bedeutung* (Sammlung Philosophie 4). Göttingen: Vandenhoeck & Ruprecht.

Ginzburg, Jonathan. 1994. An update semantics for dialogue. In Harry Bunt, Reinhard Muskens & Gerrit Rentier (eds.), *Proceedings of the first international workshop on computational semantics*. Tilburg, The Netherlands: Katholieke Universiteit Brabant.

Ginzburg, Jonathan. 1996. Dynamics and the semantics of dialogue. In Jerry Seligman & Dag Westerstøahl (eds.), *Logic, language and computation*, vol. 1 (CSLI Lecture Notes 58), 221–237. Stanford, CA: CSLI Publications.

Ginzburg, Jonathan. 1997. *On some semantic consequences of turn taking*. Paul Dekker, Martin Stokhof & Yde Venema (eds.). ILLC, University of Amsterdam.

Ginzburg, Jonathan. 2003. Disentangling public from private meaning. In Jand van Kuppevelt & Ronnie W. Smith (eds.), *Current and new directions in discourse & dialogue*, 183–211. Dordrecht: Kluwer Academic Publishers.

Ginzburg, Jonathan. 2011. Situation semantics and the ontology of natural language. In Claudia Maienborn, Klaus von Heusinger & Paul Portner (eds.), *Se-*

mantics: An international handbook of natural language meaning, vol. 1 (Handbücher zur Sprach- und Kommunikationswissenschaft), 830–851. De Gruyter Mouton. DOI:10.1515/9783110226614.830

Ginzburg, Jonathan. 2012. *The interactive stance: Meaning for conversation.* Oxford, UK: Oxford University Press.

Ginzburg, Jonathan. 2019. Acquisition. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, i–ii. Berlin: Language Science Press. DOI:??

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.

Ginzburg, Jonathan & Raquel Fernández. 2005. Scaling up from dialogue to multilogue: Some principles and benchmarks. In *Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics* (ACL '05), 231–238. Ann Arbor, Michigan: Association for Computational Linguistics. DOI:http://dx.doi.org/10.3115/12198

Ginzburg, Jonathan, Raquel Fernández & David Schlangen. 2014. Disfluencies as intra-utterance dialogue moves. *Semantics and Pragmatics* 7(9). 1–64. DOI:10.3765/sp.7.9

Ginzburg, Jonathan & Massimo Poesio. 2016. Grammar is a system that characterizes talk in interaction. *Frontiers in Psychology* 7. 1938. DOI:10.3389/fpsyg.2016.01938

Ginzburg, Jonathan & Matthew Purver. 2012. Quantification, the Reprise Content Hypothesis, and Type Theory. In Lars Borin & Staffan Larsson (eds.), *From quantification to conversation. Festschrift for Robin Cooper on the occasion of his 65th birthday.* London: College Publications.

Ginzburg, Jonathan & Ivan A. Sag. 2000. *Interrogative investigations: The form, meaning and use of English interrogatives.* Stanford: CSLI Publications.

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.

Green, Georgia M. 1996. The structure of CONTEXT: The representation of pragmatic restrictions in HPSG. In *Proceedings of the 5th annual meeting of the Formal Linguistics Society of the Midwest* (Studies in the Linguistic Sciences 24), 215–232.

Haji-Abdolhosseini, Mohammad. 2003. A constraint-based approach to information structure and prosody correspondence. In Stefan Müller (ed.), *Proceedings of the 10th International Conference on Head-Driven Phrase Structure Grammar,* *Michigan State University, East Lansing*, 143–162. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2003/, accessed 2018-2-25.

Halliday, Michael A. K. 1967. Notes on transitivity and theme in English. Part ii. *Journal of Linguistics* 3(2). 199–244.

Hamblin, Charles Leonard. 1970. *Fallacies*. London: Methuen & Co Ltd.

Heim, Irene. 2002. File change semantics. In Paul Portner & Barbara H. Partee (eds.), *Formal semantics: The essential readings* (Linguistics: The Essential Readings 1), 223–248. Oxford, UK: Wiley-Blackwell.

Heldner, Mattias & Jens Edlund. 2010. Pauses, gaps and overlaps in conversations. *Journal of Phonetics* 38(4). 555–568.

Hough, Julian, Casey Kennington, David Schlangen & Jonathan Ginzburg. 2015. Incremental semantics for dialogue processing: Requirements, and a comparison of two approaches. In *Proceedings of the 11th International Conference on Computational Semantics*, 206–216. London, UK: Association for Computational Linguistics.

Hudson, Dick. 2019. HPSG and Dependency Grammar. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, i–iv. Berlin: Language Science Press. DOI:??

Hunter, Julie & Nicholas Asher. 2015. Integrating non-linguistic events into discourse structure. In *Proceedings of the 11th international conference on computational semantics* (IWCS 2015), 184–194. London, UK.

Jackendoff, Ray. 1972. *Semantic interpretation in Generative Grammar*. Cambridge, MA: MIT Press.

Johnson, David E. & Shalom Lappin. 1999. *Local constraints vs. economy* (Stanford Monographs in Linguistics). Stanford, CA: CSLI Publications.

Kamp, Hans & Uwe Reyle. 1993. *From discourse to logic*. Dordrecht: Kluwer Academic Publishers.

Kathol, Andreas, Adam Przepiórkowski & Jesse Tseng. 2011. Advanced topics in Head-Driven Phrase Structure Grammar. In *Non-transformational syntax: Formal and explicit models of grammar*, 54–111. Oxford, UK: Wiley-Blackwell. DOI:10.1002/9781444395037.ch2

Kempson, Ruth. 2011. Formal semantics and representationalism. In Claudia Maienborn, Klaus von Heusinger & Paul Portner (eds.), *Semantics: An international handbook of natural language meaning*, vol. 1 (Handbücher zur Sprach- und Kommunikationswissenschaft / Handbooks of Linguistics and Communication Science (HSK) 33), 216–241. Berlin: De Gruyter Mouton. DOI:10.1515/9783110226614.946

Kempson, Ruth, Ronnie Cann, Eleni Gregoromichelaki & Stergios Chatzikyriakidis. 2016. Language as mechanisms for interaction. *Theoretical Linguistics* 42(3-4). 203–276. DOI:10.1515/tl-2016-0011

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

King, Jeffrey C., Scott Soames & Jeff Speaks. 2014. *New thinking about propositions*. Oxford, UK: Oxford University Press.

King, Paul J. 1999. Towards truth in Head-driven Phrase Structure Grammar. In Valia Kordoni (ed.), *Tübingen studies in head-driven phrase structure grammar*, vol. 2 (Arbeitspapiere des SFB 340 132), 301–352. Eberhard-Karls Universität Tübingen.

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. 1985. Gesprochene Sprache – geschriebene Sprache. *Zeitschrift für Literaturwissenschaft und Linguistik* 59. 9–35.

Koch, Peter & Wulf Oesterreicher. 1985. Sprache der Nähe – Sprache der Distanz. Mündlichkeit und Schriftlichkeit im Spannungsfeld von Sprachtheorie und Sprachgeschichte. In Olaf Deutschmann, Hans Flasche, Bernhard König, Margot Kruse, Walter Pabst & Wolf-Dieter Stempel (eds.), *Romanistisches jahrbuch*, 15–43. Berlin & New York: Walter de Gruyter.

Koenig, Jean-Pierre & Frank Richter. 2019. Semantics. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, i–xxxi. Berlin: Language Science Press. DOI:??

Kratzer, Angelika. 1978. *Semantik der Rede: Kontexttheorie, Modalwörter, Konditionalsätze* (Monographien: Linguistik und Kommunikationswissenschaft 38). Königstein/Ts.: Scriptor.

Krifka, Manfred. 2008. Basic notions of information structure. *Acta Linguistica Hungarica* 55(3–4). 243–276. DOI:10.1556/ALing.55.2008.3-4.2

Kripke, Saul A. 1980. *Naming and necessity*. Oxford: Blackwell.

Kubota, Yusuke. 2019. HPSG and Categorial Grammar. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, i–xlvi. Berlin: Language Science Press. DOI:??

Kuthy, Kordula De. 2019. Information structure. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, i–xxxi. Berlin: Language Science Press. DOI:??

Lee-Goldman, Russell. 2011. *Context in constructions*. Berkeley: University of California dissertation. https://escholarship.org/uc/item/02m6b3hx, accessed 2018-11-24.

Levine, Robert D. & Walt Detmar Meurers. 2006. Head-Driven Phrase Structure Grammar: Linguistic approach, formal foundations, and computational realization. In Keith Brown (ed.), *The encyclopedia of language and linguistics*, 2nd edn., 237–252. Oxford: Elsevier Science Publisher B.V. (North-Holland).

Levinson, Stephen C. & Francisco Torreira. 2015. Timing in turn-taking and its implications for processing models of language. *Frontiers in Psychology* 6(731). 10–26.

Lewis, David. 1979. Scorekeeping in a language game. In Rainer Bäuerle, Urs Egli & Arnim von Stechow (eds.), *Semantics from different points of view* (Springer Series in Language and Communication 6), 172–187. Berlin, Heidelberg & New York: Springer.

Linell, Per. 2005. *The written language bias in linguistics: Its nature, origins and transformations* (Routledge Advances in Communication and Linguistic Theory). New York: Routledge.

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. 2019. Gesture. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, i–xxxiii. Berlin: Language Science Press. DOI:??

Lücking, Andy & Jonathan Ginzburg. 2018. 'most people but not Bill': Integrating sets, individuals and negation into a cognitively plausible account of noun phrase interpretation. In *Proceedings of Cognitive Structures: Linguistic, Philosophical and Psychological Perspectives* (CoSt'18).

Lücking, Andy, Hannes Rieser & Marc Staudacher. 2006. SDRT and multi-modal situated communication. In David Schlangen & Raquel Fernández (eds.), *Proceedings of the 10th Workshop on the Semantics and Pragmatics of Dialogue* (Brandial'06), 72–79. Potsdam: Universitätsverlag Potsdam.

Marcus, R. Barcan. 1961. Modalities and intensional languages. *Synthese* 13(4). 303–322.

Martin-Löf, Per. 1984. *Intuitionistic type theory* (Studies in Proof Theory). Napoli: Bibliopolis.

McNeill, David (ed.). 2000. *Language and gesture*. Cambridge, UK: Cambridge University Press.

Montague, Richard. 1974. The proper treatment of quantification in ordinary English. In Richmond Thomason (ed.), *Formal philosophy*. New Haven: Yale UP.

Müller, Stefan. 2016. *Grammatical theory. From transformational grammar to constraint-based approaches* (Textbooks in Language Science 1). Berlin: Language Science Press. DOI:10.17169/langsci.b25.167

Müller, Stefan. 2019. HPSG and Construction Grammar. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, i–xxxix. Berlin: Language Science Press. DOI:??

Nerbonne, John. 1992. *Constraint-based semantics*. Tech. rep. 92-18. Saarbrücken: DFKI. DOI:10.22028/D291-24846

Nunberg, Geoffrey. 1978. *The pragmatics of reference*. Indiana University Bloomington, Indiana University Linguistics Club dissertation.

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

Partee, Barbara Hall. 1973. Some structural analogies between tenses and pronouns in English. *The Journal of Philosophy* 70(18). 601–609.

Penn, Gerald. 1999. An RSRL formalization of Serbo-Croatian second position clitic placement. In Valia Kordoni (ed.), *Tübingen studies in head-driven phrase structure grammar*, 177–197. Universität Tübingen.

Penn, Gerald. 2000. *The algebraic structure of attributed type signatures*. Pittsburgh, PA: School of Computer Science, Language Technologies Institute, Carnegie Mellon University dissertation.

Poesio, Massimo. 1995. A model of conversation processing based on micro conversational events. In *Proceedings of the Annual Meeting of the Cognitive Science Society*, 698–703. Pittsburgh: The Cognitive Science Society.

Poesio, Massimo & Hannes Rieser. 2010. Completions, coordination, and alignment in dialogue. *Dialogue & Discourse* 1(1). 1–89. DOI:10.5087/dad.2010.001

Poesio, Massimo & Hannes Rieser. 2011. An incremental model of anaphora and reference resolution based on resource situations. *Dialogue & Discourse* 2(1). 235–277. DOI:10.5087/dad.2011.110

Poesio, Massimo & David Traum. 1997. Conversational actions and discourse situations. *Computational Intelligence* 13(3). 309–347.

Pollard, Carl. 1997. Lectures on the foundations of HPSG. Unpublished manuscript, Ohio State University.

Pollard, Carl J. 1999. Strong generative capacity in HPSG. In Gert Webelhuth, Jean-Pierre Koenig & Andreas Kathol (eds.), *Lexical and constructional aspects of linguistic explanation*, 281–297. Stanford, CA: CSLI Publications.

Pollard, Carl & Ivan A. Sag. 1987. *Information-based syntax and semantics*. Stanford, CA: CSLI Publications.

Pollard, Carl & Ivan A. Sag. 1994. *Head-driven Phrase Structure Grammar*. Stanford, CA: CSLI Publications.

Purver, Matthew & Jonathan Ginzburg. 2004. Clarifying noun phrase semantics. *Journal of Semantics* 21(3). 283–339. DOI:10.1093/jos/21.3.283

Ranta, Aarne. 2015. Constructive type theory. In Shalom Lappin & Chris Fox (eds.), *The handbook of contemporary semantic theory*, 2nd edn., chap. 11, 343–374. Oxford, UK: Wiley-Blackwell.

Richter, Frank. 2000. *A mathematical formalism for linguistic theories with an application in head-driven phrase structure grammar*. Eberhard-Karls-Universität Tübingen dissertation.

Richter, Frank. 2004. *Foundations of lexcial ressource semantics*. Eberhard-Karls-Universität Tübingen. Habilitation.

Richter, Frank. 2019. Formal Background. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, i–ii. Berlin: Language Science Press. DOI:??

Richter, Frank & Manfred Sailer. 1999. LF conditions on expressions of Ty2: an HPSG analysis of negative concord in Polish. In Robert D. Borsley & Adam Przepiórkowski (eds.), *Slavic in HPSG*, 247–282. Stanford, CA: CSLI Publications.

Richter, Frank & Manfred Sailer. 2004. Basic concepts of Lexical Resource Semantics. In Arne Beckmann & Norbert Preining (eds.), *ESSLLI 2003 – course material I*, vol. 5 (Collegium Logicum), 87–143. Vienna: Kurt Gödel Society Wien.

Sacks, Harvey, Emanuel A. Schegloff & Gail Jefferson. 1974. A simplest systematics for the organization of turn-taking for conversation. *Language* 50(4). 696–735.

Sag, Ivan A., Thomas Wasow & Emily M. Bender. 2003. *Syntactic theory: A formal introduction*. 2nd edn. Stanford, CA: CSLI Publications.

de Saussure, Ferdinand. 2001. *Grundfragen der Allgemeinen Sprachwissenschaft*. Charles Bally & Albert Sechehaye (eds.). 3rd edn. (de Gruyter Studienbuch). Berlin: de Gruyter. Translation of *Course de linguistique générale*. Lausanne & Paris, 1916.

Schegloff, Emanuel A. & Harvey Sacks. 1973. Opening up closings. *Semiotica* 8(4). 289–327. DOI:https://doi.org/10.1515/semi.1973.8.4.289

Sedivy, Julie C., Michael K. Tanenhaus, Craig G. Chambers & Gregory N. Carlson. 1999. Achieving incremental semantic interpretation through contextual representation. *Cognition* 71(2). 109–147. DOI:http://dx.doi.org/10.1016/S0010-0277(99)00025-6

Stalnaker, Robert. 1978. Assertion. *Syntax and Semantics* 9. 315–332.

Svartvik, Jan (ed.). 1990. *The London Corpus of spoken English: Description and research* (Lund Studies in English 82). Lund: Lund University Press.

Tian, Ye, Chiara Mazzoconi & Jonathan Ginzburg. 2016. When do we laugh? In *Proceedings of the 17th Annual SIGdial Meeting on Discourse and Dialogue* (SIGDIAL 2016). Los Angeles, USA.

Tomasello, Michael. 1998. Reference: Intending that others jointly attend. *Pragmatics & Cognition* 6(1/2). 229–243. Special Issue: The Concept of Reference in the Cognitive Sciences.

Traum, David. 1994. *A computational theory of grounding in natural language conversation*. Computer Science Dept., University of Rochester dissertation.

Traum, David & Staffan Larsson. 2003. The information state approach to dialogue management. In Jan Kuppevelt, Ronnie W. Smith & Nancy Ide (eds.), *Current and new directions in discourse and dialogue* (Text, Speech and Language Technology 22), 325–353. Amsterdam: Springer.

Vallduví, Enric. 2016. Information structure. In Maria Aloni & Paul Dekker (eds.), *The Cambridge handbook of formal semantics*, 728–755. Cambridge: Cambridge University Press.

Van Eynde, Frank. 1998. Tense, aspect and negation. In Frank Van Eynde & Paul Schmidt (eds.) (Studies in Machine Translation and Natural Language Processing 10), 209–280. Office for Official Publications of the European Communities.

Van Eynde, Frank. 2000. A constraint-based semantics for tenses and temporal auxiliaries. In Ronnie Cann, Claire Grover & Philip Miller (eds.), *Grammatical interfaces in HPSG* (Studies in Constraint-Based Lexicalism), 231–249. Stanford, CA: CSLI Publications.

Wasow, Tom. 2019. Processing. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, i–xvi. Berlin: Language Science Press. DOI:??

Wechsler, Stephen. 2019. Agreement. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, i–xxi. Berlin: Language Science Press. DOI:??

Wechsler, Stephen, Jean-Pierre Koenig & Anthony Davis. 2019. Argument structure and linking. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, i–xxxvii. Berlin: Language Science Press. DOI:??

Zimmermann, Thomas Ede. 2011. Model-theoretic semantics. In Claudia Maienborn, Klaus von Heusinger & Paul Portner (eds.), *Semantics: An international handbook of natural language meaning*, vol. 1 (Handbücher zur Sprach- und Kommunikationswissenschaft / Handbooks of Linguistics and Communication Science (HSK) 33), 762–802. Berlin: De Gruyter Mouton. DOI:10.1515/9783110226614.94

Chapter 30

Sign languages

Markus Steinbach
Georg-August-Universität Göttingen

Anke Holler
Georg-August-Universität Göttingen

Change epigram in chapters/03.tex
or remove it there

Change the abstract in chapters/03.tex Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

1 Introduction

Duis pulvinar lacus id gravida ornare. Phasellus eu mauris sed tortor maximus condimentum ultrices in leo. Donec non erat nec nulla ullamcorper ornare sed id ex. Integer risus mauris, aliquet vel aliquam sed, feugiat quis nisi. Suspendisse quis nunc a turpis porttitor mollis. In luctus nulla id nunc dapibus, id rhoncus lorem pretium. Nunc eget fringilla velit, semper commodo diam. Suspendisse odio odio, euismod ac ornare sed, tincidunt ac arcu. Pellentesque vitae fringilla

Markus Steinbach & Anke Holler. N.d. Sign languages. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*. prepublished version. Berlin: Language Science Press. [Preliminary page numbering]

orci. Donec faucibus metus dui, nec iaculis purus pellentesque sit amet. Sed fermentum lorem non augue cursus, eu accumsan risus ullamcorper. Suspendisse rhoncus magna vitae enim pellentesque, eget porttitor quam finibus. Nunc ultricies turpis at quam vehicula, at tempus justo molestie. Proin convallis augue ut turpis cursus rhoncus. Donec sed convallis justo. Sed sed massa pharetra ex aliquet eleifend. finality

$$\begin{bmatrix} \textit{some-type} & \\ \textsc{feat-a} & \boxed{10} \begin{bmatrix} \textit{type-a} & \\ \textsc{feat-aa} & \textit{type-aa} \\ \textsc{feat-ab} & \left\langle \begin{bmatrix} \textsc{synsem}|\textsc{loc}|\textsc{cat}|\textsc{head} & \textit{type-aba} \\ \textsc{feat-abc} & \textit{type-abc} \end{bmatrix}, \text{NP} \right\rangle \end{bmatrix} \\ \textsc{feat-b} & \boxed{10}\,\textit{type-b} \end{bmatrix}$$

$$\begin{bmatrix} \textit{some-type} & \\ \textsc{feat-a} & \boxed{10} \begin{bmatrix} \textit{type-a} & \\ \textsc{feat-aa} & \textit{type-aa} \\ \textsc{feat-ab} & \left\langle \begin{bmatrix} \textsc{synsem}|\textsc{loc}|\textsc{cat}|\textsc{head} & \textit{type-aba} \\ \textsc{feat-abc} & \textit{type-abc} \end{bmatrix}, \text{NP} \right\rangle \end{bmatrix} \\ \textsc{feat-b} & \boxed{10}\,\textit{type-b} \end{bmatrix}$$

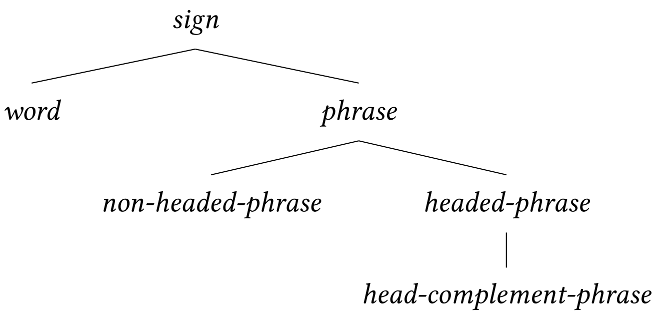


Figure 1: Type hierarchy for *sign*

Abbreviations

Acknowledgements

Chapter 31

Gesture

Andy Lücking
Université de Paris, 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 its content (Kendon 1980; McNeill 1992). The relationships between gesture and speech obey regularities that need to be defined in terms of not just the relative timing of gesture to speech, but also the linguistic form of that speech: for instance, prosody and syntactic constituency and headedness (Loehr 2007; Ebert et al. 2011; Alahverdzhieva et al. 2017). Consequently, speech-gesture integration is 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) and Abner et al. (2015), this chapter aims at distinguishing itself by providing a guided tour of research that focuses on using (mostly) standard methods for semantic composition in constraint-based grammars like HPSG to model gesture meanings.

1 Why gestures?

People talk with their whole body. A verbal utterance is couched in an intonation pattern that, via prosody, articulation speed or stress, function as *paralinguistic* signals (e.g. Birdwhistell 1970). The temporal dimension of paralinguistics

Andy Lücking. N.d. Gesture. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*. prepublished version. Berlin: Language Science Press. [Preliminary page numbering]

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* (i.e. the mutual overlap or relative timing of verbal vs. non-verbal communicative actions) is a feature of the multimodal utterance itself; it contributes, for instance, to identifying the word(s) that are affiliated to a gesture (Wiltshire 2007). A special chronemic case is signalling at the right moment – or, for that matter, 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 a more narrow sense which is also pursued from here on). 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 and contribute to discourse. 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 (see Section 4.4 for a brief discussion). Given gestures' imagistic and often spontaneous character, it is appealing to think of them as "postcards from the mind" (de Ruiter 2007: 21). Clearly, given this entrenchment in speaking, the fact that one can communicate meaning with non-verbal signals has 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.

Figure 1: *Die Skulptur die hat 'n* [*BETONsockel*] ('The sculpture has a [CONCRETE base]') [V5, 0:39]

The gesture shown in Figure 1 depicts the shape of a concrete base, which the speaker introduces into discourse as an attribute of a sculpture:[1]

(2) Die Skulptur die hat 'n [BETONsockel].
The sculpture it has a [CONCRETE base].
'The sculpture has a concrete base.'

The following representational conventions obtain: square brackets roughly indicate the portion of speech which overlaps temporally with the gesture (or more precisely, with the gesture stroke; see Figure 5 below) and upper case is used to mark main stress or accent. So both timing and intonation give clues that the gesture is related to the noun *Betonsockel* 'concrete base'. From the gesture, but not from speech, we get that the concrete base of the sculpture has the shape of a flat cylinder – thus, the gesture acts as a nominal modifier. There is a further complication, however: the gesture is incomplete with regard to its interpretation – it just depicts about half of cylinder. Thus, gesture interpretation may involve processes known from gestalt theory (see Lücking 2016 on a *good continuation* constraint relevant to (2)/Figure 1).

The speaker of the datum in Figure 2 uses just a demonstrative adverb in order to describe the shape of a building he is talking about:

[1]The examples in Figures 1, 2, 3, 4, 9 and 11 are drawn from the (German) *Speech and Gesture Alignment* corpus (SaGA, Lücking et al. 2010) and are quoted according to the number of the dialogue they appear in and their starting time in the respective video file (e.g. "V9, 5:16" means that the datum can be found in the video file of dialogue V9 at minute 5:16). Examples/Figures 4 and 9 have been produced especially for this volume; all others have also been used in Lücking (2013) and/or Lücking (2016).

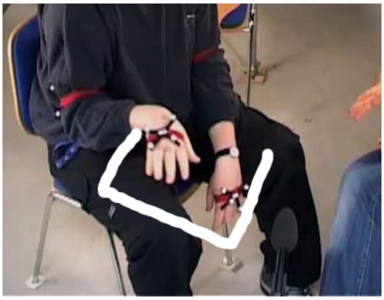


Figure 2: *Dann ist das Haus halt SO* [] ('The house is like THIS []') [V11, 2:32]

(3) Dann ist das Haus halt SO [].
Then is the house just like THIS [].
'The house is like this [].'

The demonstrative shifts the addressee's attention to the gesture, which accomplishes the full shape description, namely a cornered U-shape. In contrast to the example in Figure 1, the utterance associated with Figure 2 is not even interpretable without the gesture.

A lack of interpretability is shared by exophorically used demonstratives, which are *incomplete* without a demonstration act like a pointing gesture (Kaplan 1989: 490). For instance, Claudius would experience difficulties in understanding how serious Polonius is about his (Polonius') conjecture about the reason of Hamlet's (alleged) madness, if Polonius had not produced pointing gestures (Shakespeare, *Hamlet, Prince of Denmark* Act II, Scene 2; the third occurrence of *this* is anaphoric and refers back to Polonius' conjecture):

(4) POLONIUS (*points to his head and shoulders*): Take this from this if this be otherwise.

In order for Claudius to interpret Polonius' multimodal utterance properly, he has to associate correctly the two pointing gestures with the first two occurrences of *this* (cf. the problems discussed by Kupffer 2014). Polonius facilitates such an interpretation by means of a temporal coupling of pointing gestures and their associated demonstratives – a relationship that is called *affiliation*. The role of synchrony in multimodal utterances is further illustrated by the following example, (5), and Figure 3 (taken from Lücking 2013: 189):

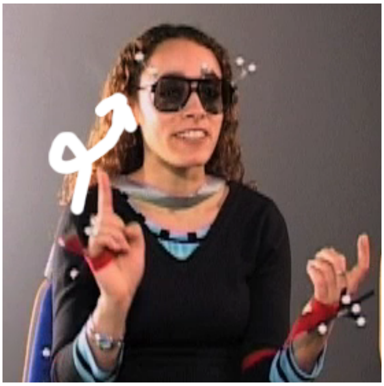


Figure 3: *Ich g[laube das sollen TREP]pen sein* (‘I think those should be STAIRcases’) [V10, 3:19]

(5) 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 between speech and gesture stroke, which is shown in Figure 3. The gesture attributes a property 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 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 a 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 3 is produced in company to stressed *glaube* (*think*) instead of *staircases*:

(6) Ich G[LAUbe das sollen Trep]pen sein.
I THINK those should staircases be
‘I think those should be staircases.’

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 temporal vicinity, prosody and syntactic constituency of the candidate affiliate (Alahverdzhieva et al. 2017).

The crucial observations in any case are that gestures contribute to propositional content and take part in pragmatic processes. Interestingly, gestures share the latter aspect with laughter, which also has propositional content (Ginzburg et al. 2015), for instance, when referring to real world events. Thus, a multimodal utterance may express a richer content than speech alone, as in (5), or a content equivalent to speech, as in (6); it can even express less than speech or contradict speech:[2]

> 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)

Contradictions or speech-gesture mismatches can occur when saying "right" but pointing left (as can be observed in everyday life but also been found in SaGA, e.g. in dialogue V24, at 4:50). A more complex case is given in (7) and Figure 4, where the speaker talks about a "rectangular arch" (which is of course a *contradictio in adiecto* in itself), but produces a roundish movement with the extended index finger of her right hand (the object she talks about is an archway). Note that the gesture just overlaps with "rectangular": its temporal extension in (7) is again indicated by means of square brackets within the original German utterances. The main stress is on the first syllable of the adjective and the noun receives secondary stress. The dots ("..") mark a short pause, so the gesture starts before "rechteckiger".

(7)	so'n	so'ne Art	[.. RECHTecki]ger	BOgen
	such an	such kind of	.. RECTangular	ARrch
	'kind of rectangular arch'			

An obvious interpretation of this mismatch is that "rectangular" is a slip of the tongue; interestingly, we found no "slip of the hand" in our data so far (which may be a hint to a possibly imagistic origin of gestures, as assumed in some production models; cf. Section 4.4).

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

Figure 4: "so'n so'ne Art [.. RECHTecki]ger BOgen" (*kind of RECTangular ARch*) [V4, 1:47].

Moving from sentence to dialogue, *interactive gestures* are bound up with turn management, among other things (Bavelas et al. 1992; 1995). For instance, pointing gestures can be used to indicate the next speaker (Rieser & Poesio 2009). Interestingly, speaker-indicating pointings are typically not produced with an outstretched index finger, but with an open hand (an example is given in Figure 14 in Section 3.6). Thus, irrespective of the question whether grammar is inherently multimodal, dialogue theory has to deal at least with certain non-verbal interaction means in any case (see also Lücking, Ginzburg & Cooper (2019), Chapter 29 of this volume).

While there is ample evidence that at least some gestures contribute to the content of the utterance they co-occur with, does this also mean that they are part of the content *intended to be communicated*? A prominent counter-example 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 be taken to be a constituent of the content intended for communication. Rather, "telephone gestures" seem to be speaker-oriented, presumably facilitating word retrieval. The fact that it is difficult to suppress gesturing even in absence of an addressee speaks in favour of a multimodal nature if not of language, then at least of speaking and surely interacting. Furthermore, the lion's share of everyday gestures seems to consist of rather sloppy movements that do not contribute to the content of the utterance in any interesting sense, though they might signal other information like speaker states. In this sense they are contingent, as opposed to being an obligatory semantic component (Lücking 2013). Gestures (or some other demonstration act) can become obligatory when they are produced within the scope of a demonstrative expression (recall (3)/Figure 2). A concurrent use with demonstratives is also one of the hallmarks collected by Cooperrider (2017) in order to distinguish *foreground* from *background* gestures (the other hallmarks are absence of speech, co-organization with speaker gaze and speaker effort). This distinction reflects two traditions within gesture studies: according to one tradition most prominently bound up with the work of McNeill (1992), gesture is a *byproduct* of speaking and therefore opens a "window into the speaker's mind". The other tradition, represented early on by Goodwin (2003) and Clark (1996), conceives gestures as a *product* of speaking, that is, as interaction means designed with a communicative intention. Since a gesture cannot be both a byproduct and a product at the same time, as noted by Cooperrider (2017), a bifurcation that is rooted in the cause and the production process of the gesture has to be acknowledged (e.g. gesturing on the phone is only puzzling from the product view, but not from the byproduct one). We will encounter this distinction again when briefly reviewing speech-gesture production models in Section 4.4. Gestures of both species are covered in the following.

2 Kinds of gestures

Pointing at an object seems to be a different kind of gesture than mimicking drinking by moving a bent hand (i.e. virtually holding something) towards the mouth while slightly rotating the back of hand upwards. And both seem to be different from actions like scratching or nose-picking. On such grounds, gestures are usually assigned to one or more classes of a taxonomy of gesture classes. Gestures that fulfil a physiological need (such as scratching, nose-picking, foot-shaking or pen-fiddling) have been called *adaptors* (Ekman & Friesen 1969) and are not dealt with further here (but see Żywiczyński et al. 2017 for evidence that adaptors may be associated with turn transition points in dialogue). Gestures that have an intrinsic relation to speech and what is communicated have been called *regulators* and *illustrators* (Ekman & Friesen 1969) and cover a variety 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. A classic taxonomy consists of the following inventory (McNeill 1992):

- iconic (or representational) gestures. Spontaneous hand and arm movements that are commonly said to be based on some kind of resemblance relation.[3] Iconic gestures employ a mode of representation such as *draw-*

[3]But see footnote 8 in Section 3.5 for pointers to critical discussions of resemblance as a sign-bearing relation.

ing, *modelling*, *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 (Wilkins 2003). In principle, any extended body part, artefact or locomotor momentum will serve the demonstrative purpose. Accordingly, there are deictic systems that involve lip-pointing (Enfield 2001) and 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). Furthermore, pointing with the hand can be accomplished by various hand shapes: Kendon & Versante (2003) distinguish *index finger pointing*, (with a *palm down* and a *palm vertical* variant) *thumb pointing*, and *open hand pointing* (again with various palm orientations). Kendon & Versante (2003: 109) claim that "the form of pointing adopted provides information about how the speaker wishes the object being indicated to be regarded". For instance, pointing with the thumb is usually used when the precise location of the intended referent is not important (Kendon & Versante 2003: 121–125), while the typical use of index finger palm down pointing is to single out an object (Kendon & Versante 2003: 115). Open hand pointing has a built-in metonymic function since the object pointed at is introduced as an example for issues related to the current discourse topic (what in semantic parlance can be conceived as the *question under discussion*; see, e.g. Ginzburg 2012). For instance, with 'open hand palm vertical', one indicates the *type* of the object pointed at instead of the object itself (Kendon & Versante 2003: 126).

- beats (rhythmic gestures, baton). 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 for an emphasising effect. "The typical beat is a simple flick of the hand or fingers up and down, or back and forth" (McNeill 1992: 15). Hence, a beat is a gestural means to accomplish what is usually expressed by vocal stress, rhythm or speed in speech.

- 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[4]). Emblems may also be more local and collected within a dictionary like the dictionary of everyday gestures in Bulgaria (Kolarova 2011).

Reconsidering gestures that have been classified as beats, among other gestures, Bavelas et al. (1992) observed that many of the stroke movements accomplish functions beyond rhythmic structuring or emphasis. Rather, they appear to contribute to dialogue management and have been called *interactive gestures.* Therefore, these gestures should be added to the taxonomy:

- interactive gestures. Hand and arm movements that accomplish the function "of helping the interlocutors coordinate their dialogue" (Bavelas et al. 1995: 394). Interactive gestures include pointing gestures that serve turn allocation ("go ahead, it's your turn") and gestures that are bound up with speaker attitudes or the relationship between speaker and addressee. Examples can be found in 'open palm/palm upwards' gestures used to indicate the information status of a proposition ("as you know") or the mimicking of quotation marks in order to signal a report of direct speech (although this also has a clear iconic aspect).

The gesture classes should not be considered as mutually exclusive categories, but rather as dimensions according to which gestures can be defined, allowing for multi-dimensional cross-classifications (McNeill 2005; Gerwing & Bavelas 2004). For instance, it is possible to superimpose pointing gestures with iconic traits. This has been found in the study on pointing gestures described in Kranstedt et al. (2006a), where two participants at a time were involved in an identification game: one participant pointed at one of several parts of a toy airplane scattered over a table, the other participant had to identify the pointed object. When pointing at a disk (a wheel of the toy airplane), some participants used index palm down pointing, but additionally turned around their index finger in a circle – that is, the pointing gesture not only locates the disk (deictic dimension) but also depicted its shape (iconic dimension). See Özyürek (2012) for an overview of various gesture classification schemes.

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

[4] 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.

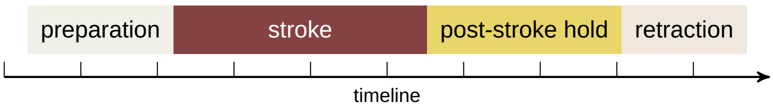


Figure 5: Gesture phases

abstract objects or scenes. For instance, an iconic drawing gesture can metaphorically display the notion "genre" via a conduit metaphor (McNeill 1992: 14):

(8) It [was a Sylves]ter and Tweety cartoon.
both hands rise up with open palm handshape, palms facing; brackets indicate segments concurrent with the gesture stroke (see Figure 5).

The gesture in (8) virtually holds an object, thus depicting the abstract concept of the genre of being a Sylvester and Tweety cartoon as a bounded container. Accordingly, gestures can be cross-classified into *concrete* and *abstract* or *metaphorical* ones (see the volume of Cienki & Müller 2008 on gesture and metaphor).

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 gesture phases are shown in the diagram in Figure 5. The stroke is the movement part that carries the gesture's meaning. It can be "frozen", leading to a post-stroke hold. If a stroke has to wait for its affiliated expression(s), a pre-stroke hold can also arise (Kita et al. 1998). The preparation and retraction phases bring hand and arms into and out of the stroke, respectively. Unless stated otherwise, when talking about gestures in what follows (and in hindsight concerning the examples given in Section 1), the stroke phase, which is the "gesture proper" or the "semantically interpretable" phase, is referred to.

Perhaps it should be noted that the spontaneous, usually co-verbal hand and arm movements considered in this chapter are different from the signed signs of sign languages (see Steinbach & Holler 2019, Chapter 30 of this volume) and pantomime (neither spontaneous nor co-verbal).[5]

3 Gestures in HPSG

Integrating a gesture's contribution into speech was initiated in computer science (Bolt 1980). Coincidentally, these early works used typed feature structure

[5]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*.

descriptions akin to the descriptive format used in HPSG grammars. Though linguistically limited, the crucial invention has been a *multimodal chart parser*, that is, an extension of chart parsing that allows the processing of input in two modalities (namely speech and gesture). Such approaches are reviewed in Section 3.2. Afterwards, a more elaborate gesture representation format is introduced that makes it possible to encode the observable form of a gesture in terms of kinematically derived attribute-value structures (Section 3.3). Following the basic semiotic distinction between deictic (or indicating or pointing) gestures and iconic (or representational or imagistic) gestures, the analysis of each class of gestures is exemplified in Sections 3.4 and 3.5, respectively. To begin with, however, some basic phenomena that should be covered by a multimodal grammar are briefly summarized in Section 3.1.

3.1 Basic empirical phenomena of grammatical gesture integration

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

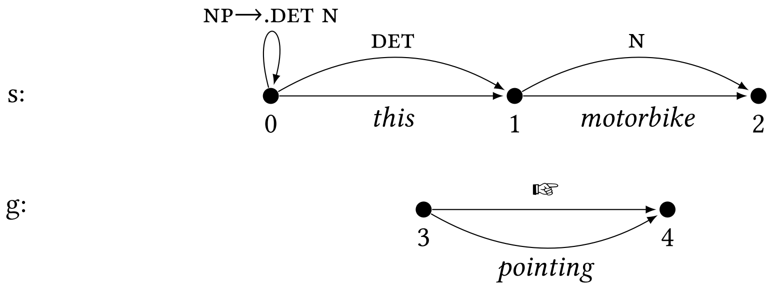
- What is the meaning of a gesture? On which grounds should semantic representations or truth conditions be assigned to hand and arm movements?
- What is the affiliate of a gesture, that is, its verbal attachment site?
- What is the result of multimodal integration, that is, the outcome of composing verbal and non-verbal meanings?

Given the linguistic significance of gestures as sketched in the preceding sections, formal grammar- and semantic-oriented accounts of speech-gesture integration have recently been developed that try to deal with (at least one of) the three basic phenomena, though with different priorities, including Alahverdzhieva (2013), Alahverdzhieva & Lascarides (2010), Ebert (2014), Giorgolo (2010), Giorgolo & Asudeh (2011), Lücking (2013; 2016), Rieser (2008; 2011; 2015), Rieser & Poesio (2009) and Schlenker (2018). It should be noted that the first basic question does not have to be considered a question for grammar, but can be delegated to a foundational theory of gesture meaning. Here gestures turn out to be like words again, where "semantic theory" can refer to explaining meaning (foundational) or specifying meaning (descriptive) (Lewis 1970: 19). In any case, the HPSG-related approaches are briefly reviewed below.

3.2 Precursors

Using typed feature structure descriptions to represent the form and meaning of gestures goes back to computer science approaches to human-computer interaction. For instance, the *QuickSet* system (Cohen et al. 1997) allows users to operate on a map and move objects or lay out barbed 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 conventional constrained-based grammar for speech-only input rests on a "unimodal" parser, Johnston (1998) and Johnston et al. (1997) developed a *multimodal chart parser*, which is still a topic of computational linguistics (Alahverdzhieva et al. 2012) (see also Bender & Emerson 2019, Chapter 28 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 a speech (s) and a gesture (g) layer:

(9)


A multimodal chart or *multichart* is defined in terms of sets of identifiers from both layers. Possible multicharts from (9) include the following ones:

(10) 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 (11):

(11)
$$
\begin{bmatrix}
\textsc{lhs}: & \begin{bmatrix} \textsc{cat}: & \textit{command} \\ \textsc{modality}: & \boxed{2} \\ \textsc{content}: & \boxed{1} \\ \textsc{time}: & \boxed{3} \end{bmatrix} \\
\textsc{rhs}: & \begin{bmatrix} \textsc{dtr1}: & \begin{bmatrix} \textsc{cat}: & \textit{located\_command} \\ \textsc{modality}: & \boxed{6} \\ \textsc{content}: & \boxed{1}[\textit{location}\ \boxed{5}] \\ \textsc{time}: & \boxed{7} \end{bmatrix} \\ \textsc{dtr2}: & \begin{bmatrix} \textsc{cat}: & \textit{spatial\_gesture} \\ \textsc{content}: & \boxed{5} \\ \textsc{modality}: & \boxed{9} \\ \textsc{time}: & \boxed{10} \end{bmatrix} \end{bmatrix} \\
\textsc{constraints}: & \left\{ \begin{array}{l} \textsc{overlap}(\boxed{7},\boxed{10}) \vee \textsc{follow}(\boxed{7},\boxed{10},4\textsc{s}) \\ \textsc{total-time}(\boxed{7},\boxed{10},\boxed{3}) \\ \textsc{assign-modality}(\boxed{6},\boxed{9},\boxed{2}) \end{array} \right\}
\end{bmatrix}
$$

The AVM in (11) 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 (see the 4s under CONSTRAINTS). An integration scheme highly akin to that displayed in (11) also underlies current grammar-oriented approaches to deictic and iconic gestures (see Sections 3.4 and 3.5 below).

3.3 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 3 according to the format used in Lücking et al. (2010) is given in (12):

(12)
$$\begin{bmatrix} \textit{right hand} & \\ \textsc{handshape} & \begin{bmatrix}\textsc{shape} & \textit{G}\\ \textsc{path} & \textit{0} \\ \textsc{dir} & \textit{0}\end{bmatrix} \\ \textsc{palm} & \begin{bmatrix}\textsc{orient} & \textit{PAB>PAB/PUP>PAB}\\ \textsc{path} & \textit{0} \\ \textsc{dir} & \textit{0}\end{bmatrix} \\ \textsc{boh} & \begin{bmatrix}\textsc{orient} & \textit{BUP>BTB/BUP>BUP}\\ \textsc{path} & \textit{arc>arc>arc} \\ \textsc{dir} & \textit{MR>MF>ML}\end{bmatrix} \\ \textsc{wrist} & \begin{bmatrix}\textsc{position} & \textit{P-R}\\ \textsc{path} & \textit{line} \\ \textsc{dir} & \textit{MU} \\ \textsc{dist} & \textit{D-EK} \\ \textsc{extent} & \textit{small}\end{bmatrix} \\ \textsc{sync} & \begin{bmatrix}\textsc{config} & \textit{BHA}\\ \textsc{rel.mov} & \textit{LHH}\end{bmatrix} \end{bmatrix}$$

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 6. 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 for the whole hand 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 7. 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, also acknowledging the distance of the hand from the speaker's body (feature dist) turns the shallow disk into a three-dimensional space, giving rise to the three-dimensional model displayed on the right-hand side of Figure 7. The gesture space regions known as *center-center*, *center* and *periphery*, possibly changed by location modifiers (*upper right*, *right*, *lower right*, *upper left*, *left*, *lower left*), are now modelled as nested cuboids. Thus, gesture space is structured according to all three body axes: the sagittal, the longitudinal and the transverse axes. Annotations straightforwardly transfer to the three-dimensional gesture

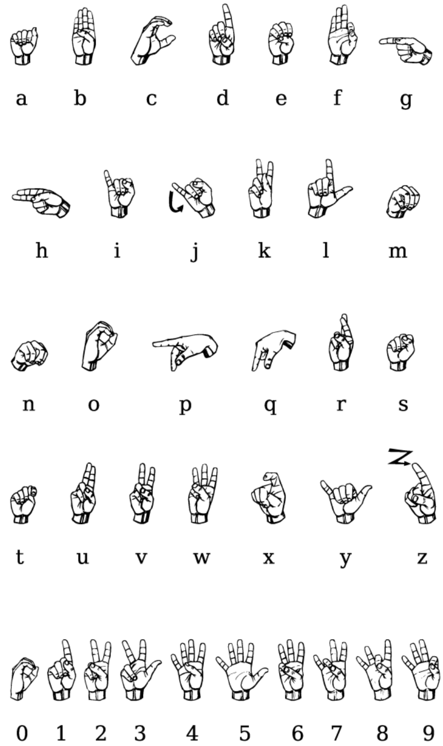


Figure 6: American Sign Language fingerspelling alphabet (image released to Public Domain by user Ds13 in the English Wikipedia at 18th December 2004, https://commons.wikimedia.org/wiki/File:Asl_alphabet_gallaudet.png)

space model. Such a three-dimensional gesture space model is assumed throughout this chapter. Complex movement trajectories through the vector space can describe a rectangular or a roundish path (or mixtures of both). Both kinds of movements are distinguished in terms of *line* or *arc* values of feature PATH. An example illustrating the difference is given in Figure 8. A brief review of gesture annotation can be found in Section 4.1.

3.4 Pointing Gestures

Pointing gestures are *the* prototypical referring device: they probably pave a way to reference in both evolutionary and language acquisition perspectives (Bruner

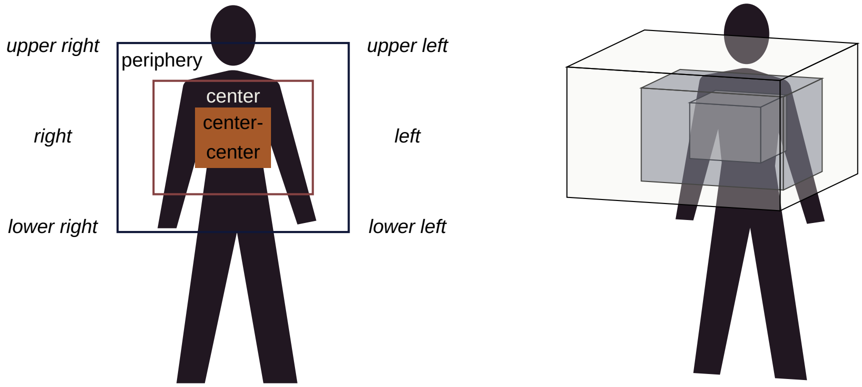


Figure 7: Gesture Space (left hand side is simplified from McNeill 1992: 89). Although originally conceived as a structured "shallow disk" McNeill (1992: 86), adding distance information gives rise to a three-dimensional gesture space model as illustrated on the right-hand side.

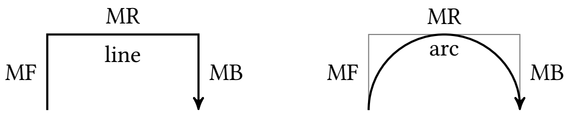


Figure 8: The same sequence of direction labels can give rise to an open rectangle or a semicircle, depending on the type of concatenation (Lücking 2016: 385).

1998; Masataka 2003; Matthews et al. 2012); they are predominant inhabitants of the "deictic level" of language, interleaving the symbolic (and the iconic) levels (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).

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, with this conception, not only utterance production but *grammar* is inherently multimodal.

The referential import of the pointing gesture has been studied experimentally in 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 rely 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* (Kranstedt et al. 2006b; Lücking et al. 2015). These works provide an answer to the first basic question (cf. Section 3.1): 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. In a related way, two different functions of a pointing gesture have been distinguished by Kühnlein et al. (2002), namely singling out an object (13a) or making an object salient (13b).

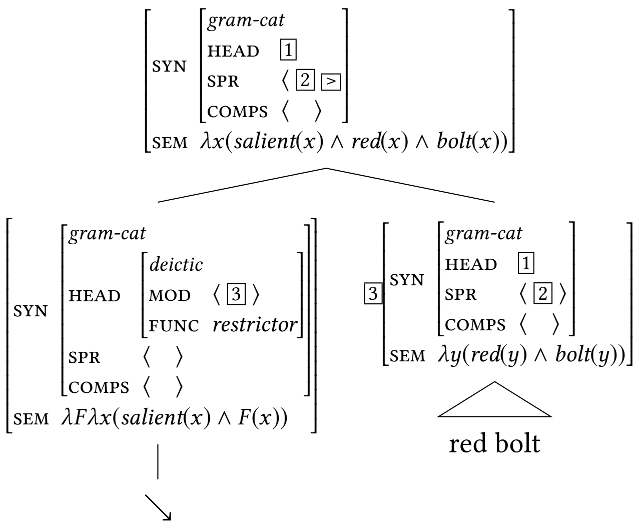
(13) a. $\lambda F \lambda x(x = c \wedge F(x))$

b. $\lambda F \lambda x(\mathit{salient}(x) \wedge F(x))$

The approach is expressed in lambda calculus and couched in an HPSG framework. The derivation of the instruction *Take the red bolt* plus a pointing gesture is exemplified in (14). A pointing gesture is represented by means of "↘" and takes a syntactic position within the linearized inputs according to the start of the stroke phase. For instance, the pointing gesture in (14a) occurred after *the* has been articulated but before *red* is finished. The derivation of the multimodal N′ constituent is shown in (14b).

(14) a. Take [the ↘[N′ [N′ red bolt]]]

b.



The spatial model is also adopted in Lascarides & Stone (2009), where the region denoted by pointing is represented by a vector $\vec{p}$. This region is an argument

Figure 9: *Und man*[*chmal ist da auch ein EISverkäufer*] ('and some[times there's an ICE cream guy]'), [V5, 7:20]

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 the real world example in (15) and Figure 9, taken from dialogue V5 of the SaGA corpus.

(15) Und man[chmal ist da auch ein EISverkäufer].
And some[times] is there also an ICE cream guy.
'And sometimes there's an ice cream guy'

The context in which the gesture appears is the following: the speaker describes a route which goes around a pond. He models the pond with his left hand, a post-stroke hold (cf. Figure 5) held over several turns. After having drawn the route around the pond with his right hand, the pointing gesture in Figure 9 is produced. The pointing indicates the location of an ice cream vendor in relation to the pond modelled in gesture space. Such instances of indirect or proxy pointing have been interpreted as *dual points* by Goodwin (2003); in standard semantics they are analysed in terms of *deferred reference*, where one thing is indicated but another but related thing is referred to (Quine 1950; Nunberg 1993). The "duality" or "deference" involved in the datum consists of a mapping from the location indicated in gesture space onto a spatial area of the described real world situation. Such mappings are accounted for by the function v that shifts the pointing cone area from gesture space $\vec{p}$ to some other space $v(\vec{p})$ (Lascarides & Stone 2009). So, the deictic gesture locates the ice cream vendor. Since it is held during nearly the whole utterance, its affiliate expression "Eisverkäufer" (*ice cream* *guy*) is picked out due to carrying primary accent (indicated by capitalization).[6] 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) – see also Kuthy (2019), Chapter 24 of this volume. The well-developed basic integration scheme of Alahverdzhieva et al. (2017: 445) rests on a strict speech and gesture overlap and is called the *Situated Prosodic Word Constraint*, which allows the combination of a speech daughter (S-DTR) and a gesture daughter (G-DTR) – see Figure 10. The Situated Prosodic Word Constraint applies to both deictic and iconic gestures. Under certain conditions, including when a deictic gesture is direct (i.e. $\vec{p} = v(\vec{p})$), however, the temporal and prosodic constraints can be relaxed for pointings.

$$
\left[\begin{array}{ll}
\multicolumn{2}{l}{\textit{word}} \\
\textsc{overlap} & \langle @7, @8 \rangle \\
\textsc{time} & @7 \cup @8 \\
\textsc{phon} & @3 \\
\textsc{synsem} & \left[\begin{array}{ll}
\textsc{cat} & @5 \\
\textsc{cont} & \left[\begin{array}{ll}
\textsc{rels} & \boxed{C_{rel}} \oplus \boxed{S_{rel}} \oplus \boxed{G_{rel}} \\
\textsc{hcons} & \boxed{S_{hc}} \oplus \boxed{G_{hc}}
\end{array}\right]
\end{array}\right] \\
\textsc{s-dtr} & \left[\begin{array}{ll}
\multicolumn{2}{l}{\textit{word}} \\
\textsc{time} & @7 \\
\textsc{phon} & @3\,\textit{nuclear\_or\_pre-nuclear} \\
\textsc{synsem} & \left[\begin{array}{ll}
\textsc{cat} & @5 \\
\textsc{cont} & \left[\begin{array}{ll}
\textsc{rels} & \boxed{S_{rel}} \\
\textsc{hcons} & \boxed{S_{hc}}\,\textit{hcons}
\end{array}\right]
\end{array}\right]
\end{array}\right] \\
\textsc{g-dtr} & \left[\begin{array}{ll}
\multicolumn{2}{l}{\textit{depicting\_or\_deictic}} \\
\textsc{time} & @8 \\
\textsc{synsem} & \left[\begin{array}{ll}
\textsc{cat} & \left[\begin{array}{ll}
\textsc{g-feature} & \textit{value} \\
\ldots &
\end{array}\right] \\
\textsc{cont} & \left[\begin{array}{ll}
\textsc{rels} & \boxed{G_{rel}} \\
\textsc{hcons} & \boxed{G_{hc}}
\end{array}\right]
\end{array}\right]
\end{array}\right] \\
\textsc{c-cont.rels} & \boxed{C_{rel}}
\end{array}\right]
$$

Figure 10: *Situated Prosodic Word Constraint* (Alahverdzhieva et al. 2017: 445)

[6] Semantically, other integration points are possible, too, most notably with “da” (*there*). However, the intonation-based integration point patterns well with observations of the affiliation behaviour of iconic gestures, as indicated with respect to examples (5) and (6) in Section 1. Concerning deictic gestures, a constraint that favours affiliation to deictic words over affiliation to stressed words (if they differ at all) seems conceivable nonetheless.

In order to deal with gestures that are affiliated with expressions that are larger than single words, Alahverdzhieva et al. (2017) 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). In this account, the affiliation problem (the second desideratum identified in Section 3.1) has a well-motivated solution on both the word and the phrasal levels, at least for temporally overlapping speech-gesture occurrences (modulo the conditioned relaxations for pointings). Semantic integration of gesture location and verbal meaning (the third basic question from Section 3.1) is brought about using the underspecification mechanism of *Robust Minimal Recursion Semantics* (RMRS), a refinement of *Minimal Recursion Semantics* (MRS) (Copestake et al. 2005), where basically scope as well as arity of elementary expressions is underspecified (Copestake 2007) – see the RELS and HCONS features in Figure 10. For some background on (R)MRS see the above given references, or see Koenig & Richter (2019), Chapter 23 of this volume.

A dialogue-oriented focus on pointing is taken in Lücking (2018): here, pointing gestures play a role in formulating processing instructions that guide the addressee in where to look for the referent of demonstrative noun phrases.

3.5 Iconic Gestures

There is nearly no semantic work on the grounds according to which the meanings assigned to iconic gestures should be assigned to them in the first place (this is the first basic question from Section 3.1). Semantic modelling usually focuses on the interplay of (in this sense presumed) gesture content with speech content, that is, on the third of the basic questions from Section 3.1. Schlenker (2018: 296) is explicit in this respect: "It should be emphasized that we will not seek to explain how a gesture [...] comes to have the content that it does, but just ask how this content interacts with the logical structure of a sentence".[7] Two exceptions, however, can be found in 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. Thus, he tries to address the first basic question from Section 3.1 in terms of an empirically extracted gesture typology. These geometrical objects are used in order to provide a possibly underspecified semantic representation for iconic gestures, which is then integrated into word meaning via lambda calculus (Hahn & Rieser 2010; Rieser 2011). The work of Lücking (2013;

[7] The omission indicated by "[...]" just contains a reference to an example in the quoted paper.

2016) is inspired by Goodman's notion of *exemplification* (Goodman 1976), that is, 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 order to make this approach work, common perceptual features for predicates are extracted from their denotation and represented as part of a lexical extension of their lexemes, serving as an interface between hand and arm movements and word meanings. This conception in turn is motivated by psychophysic theories of the perception of biological events (Johansson 1973), draws on philosophical similarity conceptions beyond isomorphic mappings (Peacocke 1987),[8] and, using a somewhat related approach, has been proven to work in robotics by means of imagistic description trees (Sowa 2006). These perceptual features serve as the integration locus for iconic gestures, using standard unification techniques. The integration scheme for achieving this is the following one (Lücking 2013: 249) (omitting the time constraint used in the basic integration scheme in 11):

(16)
$$\left[\begin{array}{ll}
\multicolumn{2}{l}{\textit{sg-ensemble}} \\
\textsc{phon} & \boxed{12} \\
\textsc{cat} & \boxed{2} \\
\textsc{cont} & \boxed{3}\left[\textsc{restr}\ \left\langle \ldots, \boxed{5}\left[\begin{array}{ll} \multicolumn{2}{l}{\textit{pred}} \\ \textsc{cvm} & \boxed{1} \end{array}\right], \ldots \right\rangle\right] \\
\textsc{s-dtr} & \left[\begin{array}{ll} \multicolumn{2}{l}{\textit{verbal-sign}} \\ \textsc{phon} & \boxed{12}\left[\textsc{accent}\ \boxed{6}\right] \\ \textsc{cat} & \boxed{2} \\ \textsc{cont} & \boxed{3} \end{array}\right] \\
\textsc{g-dtr} & \left[\begin{array}{ll} \multicolumn{2}{l}{\textit{gesture-vec}} \\ \textsc{aff} & \left\langle\left[\textsc{phon}\ \left[\textsc{accent}\ \boxed{6}\textit{marked}\right]\right]\right\rangle \\ \textsc{traj} & \boxed{1} \\ \textsc{cont} & \left[\begin{array}{ll} \textsc{mode} & \textit{exemplification} \\ \textsc{ex-pred} & \boxed{5} \end{array}\right] \end{array}\right]
\end{array}\right]$$

Comparable to a modifier, a gesture attaches to an affiliate via feature AFF, which in turn is required to carry intonational accent, expressed in terms of information packaging developed by Engdahl & Vallduví (1996) (cf. Kuthy 2019, Chapter 24 of this volume). The semantic contribution of a gesture is contributed via the new semantic mode *exemplification*, that is, a gesture displays a predication from the RESTR list of its affiliate. The exemplification interface is established

[8]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).

using the format of vector semantics developed by Zwarts & Winter (2000) and Zwarts (2003) in order to capture the semantic contribution of locative prepositions, motion verbs and shape adjectives, among other things. This involves two steps: on the one hand, the representation of a gesture (cf. Section 3.3) is mapped onto a vectorial representation; on the other hand, the content of place and form predicates is enriched by abstract psychophysic information in the sense of Johansson (1973) (see above), also spelled out in terms of vector representations. Both steps are illustrated by means of the simple example shown in Figure 11, where the speaker produces a semicircle in both speech and gesture.

Figure 11: “und [oben haben die so’n HALBkreis]” (*and on the top they have such a SEMIcircle*), [V20, 6:36].

The kinematic gesture representation of the movement carried out (CARRIER) by the wrist – *move up*, *move left*, *move down*, which are concatenated (“$\oplus$”) by movement steps in a bent (“$\oplus_{\smile}$”, as opposed to rectangular “$\oplus_{\llcorner}$”) way (cf. also Figure 8) – is translated via vectorising function **V** into a vector trajectory (TRAJ(ECTORY)) from the three-dimensional vector space, cf. Figure 7:[9]

(17)
$$\begin{bmatrix} \textit{gesture-vec} & \\ \text{TRAJ} & \left[\mathbf{V}\boxed{1}=\mathbf{UP} \oplus_{\smile} -\mathbf{RT} \oplus_{\smile} -\mathbf{UP}\right] \\ \text{CARRIER} & \begin{bmatrix} \textit{gesture} & \\ \text{MORPH} & \left[\text{WRIST.MOV}\ \boxed{1}\,\textit{mu} \oplus_{\smile} \textit{ml} \oplus_{\smile} \textit{md}\right] \end{bmatrix} \end{bmatrix}$$

The lexical entry for *semicircle* is endowed with a *conceptual vector meaning* attribute CVM. Within CVM it is specified (or underspecified) what kind of vector (VEC) is at stake (axis vector, shape vector, place vector), and how it looks, that

[9] Vectors within gesture space can be conceived of as equivalence classes over concrete movement annotation predicates.

is, which PATH it describes. A semicircle can be defined as an axis vector whose path is a 180° trajectory. Accordingly, 180° is the root of a type hierarchy which hosts all vector sequences within gesture space that describe a half circle. This information is added in terms of a form predicate to the restriction list of *semicircle*, as shown in the speech daughter's (S-DTR) content (CONT) value in (18). Licensed by the speech-gesture integration scheme in (16), the half-circular gesture trajectory from (17) and its affiliate expression *semicircle* can enter into an ensemble construction, as shown in (18):

(18)
$$
\left[\begin{array}{ll}
\multicolumn{2}{l}{\textit{sg-ensemble}} \\
\textsc{phon} & \left\langle \boxed{1}\left[\begin{array}{ll}\textsc{seg} & \textit{semicircle}\end{array}\right]\right\rangle \\
\textsc{cat} & \boxed{2}\left[\begin{array}{ll}\textsc{head} & \textit{noun}\end{array}\right] \\
\textsc{cont} & \boxed{3}\left[\begin{array}{ll}
\textsc{index} & i \\
\textsc{restr} & \left\langle \left[\begin{array}{ll}\multicolumn{2}{l}{\textit{geom-obj-pred}} \\ \textsc{reln} & \textit{semicircle} \\ \textsc{inst} & i\end{array}\right], \boxed{5}\left[\begin{array}{ll}\multicolumn{2}{l}{\textit{form-pred}} \\ \textsc{reln} & \textit{round2} \\ \textsc{arg} & i \\ \textsc{cvm} & \left[\begin{array}{ll}\textsc{vec} & \textit{axis-path}(i,\boldsymbol{v}) \\ \textsc{path} & \boxed{7}\end{array}\right]\end{array}\right]\right\rangle
\end{array}\right] \\
\textsc{s-dtr} & \left[\begin{array}{ll}\multicolumn{2}{l}{\textit{verbal-sign}} \\ \textsc{phon} & \boxed{1}\left[\begin{array}{ll}\textsc{accent} & \boxed{6}\end{array}\right] \\ \textsc{cat} & \boxed{2} \\ \textsc{cont} & \boxed{3}\end{array}\right] \\
\textsc{g-dtr} & \left[\begin{array}{ll}\multicolumn{2}{l}{\textit{gesture-vec}} \\ \textsc{aff} & \left\langle\left[\begin{array}{ll}\textsc{phon} & \left[\begin{array}{ll}\textsc{accent} & \boxed{6}\textit{marked}\end{array}\right]\end{array}\right]\right\rangle \\ \textsc{traj} & \boxed{7}\boldsymbol{UP} \oplus_{\smile} -\boldsymbol{RT} \oplus_{\smile} -\boldsymbol{UP} \\ \textsc{cont} & \left[\begin{array}{ll}\textsc{mode} & \textit{exemplification} \\ \textsc{ex-pred} & \boxed{5}\end{array}\right]\end{array}\right]
\end{array}\right]
$$

By extending lexical entries with frame information from frame semantics (Fillmore 1982), also the exemplification of non-overtly-expressed predicates becomes feasible (Lücking 2013: Sec. 9.2.1); a datum showing this case has already been given with the *spiral staircases* in (5)/Figure 3. A highly improved version of the "vectorisation" of gestures with a translation protocol has been spelled out in Lücking (2016), but within the semantic framework of a *Type Theory with Records* (Cooper 2019; Cooper & Ginzburg 2015; cf. also Lücking, Ginzburg & Cooper 2019, Chapter 29 of this volume).

The richer formal, functional and representational features of iconic gestures as compared to deictic gestures (cf. Section 3.4) is accounted for in Alahverdzhieva et al. (2017) by assigning a formal predicate to each "phonological" feature of a gesture representation (cf. Section 3.3). These formal gesture predicates are highly underspecified, using *Robust Minimal Recursion Semantics* (RMRS) (Copestake 2007). That is, they can be assigned various predications (which are assumed to be constrained by iconicity with differing arity in the gesture resolution process).

Let us illustrate this by means of Example 1 from Alahverdzhieva et al. (2017), which is due to Loehr (2004) and re-given in (19), adapted to the representational conventions followed in this chapter.

(19) [So he mixes MUD]
The speaker performs a circular movement with the right hand over the upwards, open palm of the left hand

Using a variant of a kinematic representation format for gestures (cf. Section 3.3), the right hand from example 19 is notated as follows (Alahverdzhieva et al. 2017: 440):

(20)
$$\begin{bmatrix} \textit{depict-literal} & \\ \textsc{hand-shape} & \textit{bent} \\ \textsc{palm-orient} & \textit{towards-down} \\ \textsc{finger-orient} & \textit{towards-down} \\ \textsc{hand-location} & \textit{lower-periphery} \\ \textsc{hand-movement} & \textit{circular} \end{bmatrix}$$

Each feature value pair from the gesture's representation in (20) is mapped onto an RMRS-based underspecified representation (Alahverdzhieva et al. 2017: 442):

(21)
$$\begin{array}{l} l_0 : a_0 : [\mathcal{G}](h) \\ l_1 : a_1 : hand\_shape\_bent(i_1) \\ l_2 : a_2 : palm\_orient\_towards\_down(i_2) \\ l_3 : a_3 : finger\_orient\_towards\_down(i_3) \\ l_4 : a_4 : hand\_location\_lower\_periphery(i_4) \\ l_5 : a_5 : hand\_movement\_circular(i_5) \\ h =_q l_n \ \textit{where}\ 1 \leq n \leq 5 \end{array}$$

Note that all predicates mapped from the gesture in (21) fall within the scope of the scopal operator $[\mathcal{G}]$; this prevents an individual introduced by a depicting gesture from being an antecedent of a pronoun in speech.

Regimented by the *Situated Prosodic Word Constraint* from Figure 10, the underspecified semantic description of the gesture in (21) and its affiliated noun

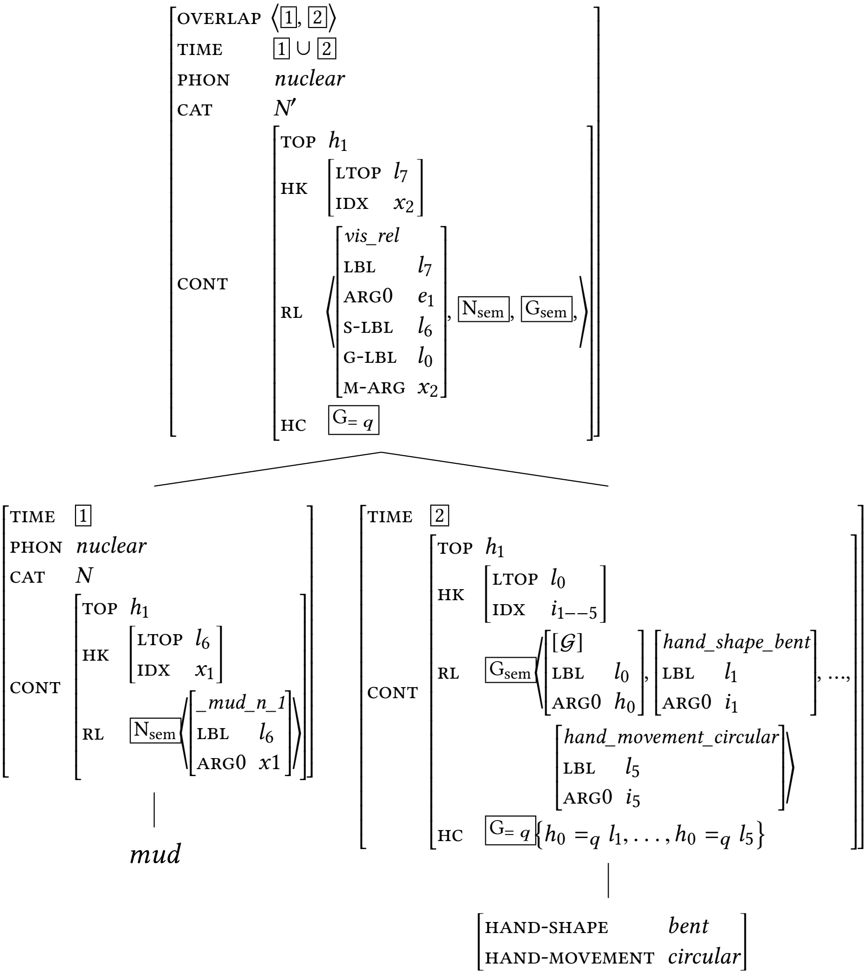


Figure 12: Derivation tree for depicting gesture and its affiliate noun *mud* (Alahverdzhieva et al. 2017: 447)

mud can enter into the multimodal construction given in Figure 12 (where the gesture features are partly omitted for the sake of brevity).

The underspecified RMRS predicates derived from gesture annotations are interpreted according to a type hierarchy rooted in those underspecified RMRS predicates. For example, the circular hand movement of the "mud gesture" can give rise to two slightly different interpretations: on the one hand, the circular hand movement can depict – in the context of the example – that mud is being mixed from an observer viewpoint (McNeill 1992). This reading is achieved by following the left branch of (22), where the gesture contributes a conjunction of predications that express that a substance rotates. When integrated with speech, the substance resolves to the mud and the rotating event to the mixing. On the other hand, the gesture can depict seen from the character viewpoint (McNeill 1992), which corresponds to the predication from the right branch of (22). Here the rotating event is brought about by agent j' which is required to be coreferential with *he*, the subject of the utterance.

(22)

hand_movement_circular(i)

substance(x') ∧ *rotate*(e', x')   rotate(e', j', x')

In addition to addressing (solving) the three basic questions identified in Section 3.1 – roughly, foundation of gesture meaning, regimenting affiliation, and characterisation of semantic integration – another issue has received attention recently, namely the projection behaviour of gestures when interacting with logical operators (Ebert 2014; Schlenker 2018). For instance, the unembedded gesture in (23) triggers the inference that the event being described actually happened in the manner in which it was gesticulated (Schlenker 2018: 303):

(23) John [*slapping gesture*] punished his son.
⇒ John punished his son by slapping him.

That is, (23) more or less corresponds to what semantic speech-gesture integration approaches, as briefly reviewed above, would derive as the content of the multimodal utterance.

Embedding the slapping gesture under the *none*-quantifier triggers, according to Schlenker (2018: 303), the following inference:

(24) None of these 10 guys [*slapping gesture*] punished his son.
⇒ for each of these 10 guys, if he had punished his son, this would have involved some slapping.

The universal inference patterns with presupposition. Unlike presupposition, however, Schlenker (2018: 303) claims that the inference is conditionalized on the at-issue contribution of (24), expressed by the *if*-clause. He then develops a notion of "cosupposition", which rests on an expression's local context that entails the content of its affiliated gesture. So far, there is no connection from such projections to HPSG, however.

Beyond being involved in pragmatic processes like inferring, gestures also take part in "micro-evolutionary" developments. 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 of sign formation as studied, for instance, in experimental semiotics (Galantucci & Garrod 2011). Such grammaticalisation processes eventually might lead to conventional signs. However, most conventional, emblematic everyday gestures seem to be the result of circumventing a taboo: something you should not name is gesticulated (Posner 2002).

3.6 Other gestures

As noted in the taxonomy reviewed in Section 2, there are gestures that, unlike the deictic and iconic ones discussed in the previous sections, do not contribute to propositional content, but serve functions bound up with dialogue management. Such gestures have been called *interactive gestures* (Bavelas et al. 1992). Two examples are given in Figures 13 and 14, which have been discussed by Bavelas et al. (1995).

The "delivery gesture" in Figure 13 is used to underline an argument, or to refer to the fact that the current issue is known to the interlocutors. In the latter function, the gesture is also termed *shared information gesture*.

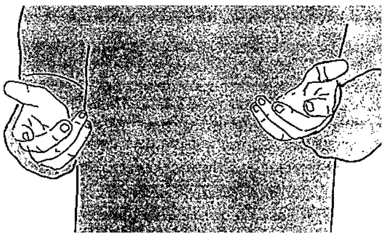


Figure 13: "Here's my point."