Elisabeth Zima* and Alexander Bergs

Multimodality and construction grammar

DOI 10.1515/lingvan-2016-1006

The meaning-making process in face-to-face interaction relies on the integration of meaningful information being conveyed by speech as well as the tone of voice, facial expressions, hand and head gestures, body postures and movements (McNeill 1992; Kendon 2004). Hence, it is inherently multimodal. Usage-based linguistics attributes language use a fundamental role in linguistic theorizing by positing that the language system is grounded in and abstracted from (multimodal) language use. However, despite this inherent epistemological link, usage-based linguists have hitherto conceptualized language as a system of interconnected *verbal*, i. e. monomodal units, leaving nonverbal usage aspects and the question of their potential entrenchment as part of language largely out of the picture.

This is – at least at first sight – surprising because the usage-based model of Construction Grammar (C × G) seems particularly well-equipped to unite the natural interest of linguists in the units that define language systems with the multimodality of language use. Constructions are conceptualized as holistic "conventionalized clusters of features (syntactic, prosodic, pragmatic, semantic, textual, etc.) that recur as further indivisible associations between form and meaning" (Fried 2015: 974). Given its conceptual openess to all levels of usage features, several studies have recently argued for the need to open up the current focus of C × G towards kinesic recurrences (Günthner & Imo 2006; Deppermann 2011; Deppermann & Proske 2015; Andrén 2010; Schoonjans 2014; Schoonjans et al. 2015; Steen & Turner 2013; Zima 2014a; Zima 2014b, in press; Cienki 2012; Cienki 2015; Mittelberg 2014; Müller & Bressem 2014; Bergs 2015; Valenzuela 2015). Departing from the usage-based foundation of C × G which takes "grammar to be the cognitive organization of one's experience with language" (Bybee 2006: 219), these studies suggest that the basic units of language, i. e. constructions, may be multimodal in nature.

This paper presents some of the current issues for a Multimodal Construction Grammar. The aim is to frame the debate and to briefly summarize some of the discussion's key issues. The individual papers in the special issue elaborate in more detail on particular points of discussion and/or present empirical case studies.

1 Seeds of multimodal construction grammar: Multimodal language learning and the usage-based hypothesis

Usage-based versions of C × G subscribe to the view that constructions are learnt through the abstraction of recurrent concrete form-meaning patterns from the linguistic input that language users are exposed to. Most notably, for children who acquire their first language(s), this input is fundamentally multimodal. Constructions are learnt in rich contextual environments and a wealth of research has shown that multimodal cues play a pivotal role in their acquisition processes. Most notably, studies that relate infants' gesture productions to their cognitive developmental stages and their communicational skills describe gesture use as facilitating children's access to first verbal combinations and argue that the type and complexity of their gestures foreshadow and even predict their verbal communicative skills (Goldin-Meadow 2013).

It is also uncontroversial that the language acquisition process involves the gradual emancipation of the verbal system from its dependency on co-speech gesture use (see also Cienki). This emancipation process, however, is not simply reflected in a decrease in frequency of gestures. Rather, it involves a gradual shift of

***Corresponding author: Elisabeth Zima**, E-mail: elisabeth.zima@germanistik.uni-freiburg.de
Alexander Bergs, E-mail: abergs@uos.de.

the semantic-pragmatic functions exerted by the verbal and the gestural modality, and the development towards one integrative multimodal language system in adult speakers (Goldin-Meadow & Butcher 2003; McNeill 1992). Within this system, gesture-speech combinations fall on a continuum with respect to the semantic overlap between the two modalities. They range from co-expressive to cases in which very disparate meanings are expressed in speech and gesture (Goldin-Meadow & Butcher 2003).

From a C × G perspective, the co-expressiveness of (a portion) of speakers' speech-gesture combinations as well as the fact that gestures are *recurrent* but usually not *obligatory* usage features, is often taken as an argument for that gestures (and probably even more so postures and facial expressions) are not part of the linguistic system, but rather non-entrenched additives of language use (cf. Ningelgen & Auer, Ziem). However, some authors argue that it is arbitrary to include optional verbal elements in constructions, but to exclude optional, yet recurrent non-verbal usage features (cf. Schoonjans).

The debate whether (some) constructions are multimodal should in any case not be confused with the rather trivial claim that constructs, i. e. instantiations, are multimodal (also argued by Ziem, Ningelgen & Auer, Cienki). Rather, to have constructional status, a given gesture-speech combination needs to be "entrenched in the minds of the speakers and conventional in the speech community" (Langacker 2008: 250 f.). Indeed, the discussion of whether to take kinesic recurrence as evidence for its entrenchment at the constructional level critically reflects the ongoing discussion in 'traditional', i. e. verbal-oriented C × G on the nature of the linguistic sign. The papers by Alan Cienki, Thomas Hoffmann, Steven Schoonjans, and Alexander Ziem all focus on the conceptual nature of the 'constructional sign' and discuss, on the one hand, prerequisites for recurrent multimodal usage patterns to qualify as constructions and, on the other hand, consequences for attempts to include multimodal information in C × G. The following section gives a summary of the main points raised in the contributions to the special issue.

2 Multimodal signs? Gestures as part of structuralist form-meaning pairings?

Constructions are abstracted from language use as conventional and entrenched associations of form and function (cf. Fried 2015). Whereas a construction's conventionality is considered to be mirrored in its acceptability with members of a speech community, entrenchment is a much fuzzier notion. This is partly due to considerable changes in the definition of constructions in cognitive versions of C × G (for discussion of entrenchment in usage-based linguistics see among others Schmid 2007; Blumenthal-Dramé 2012). In early versions of C × G, the term was confined to syntactically or semantic-pragmatically unpredictable, idiosyncratic, yet conventional linguistic structures (cf. e. g. Fillmore et al. (1988) seminal analyses of [the Xer, the Y-er] and [let alone]). Acknowledging the role of frequency in unit formation and processing, Adele Goldberg, however, has proposed an extended definition of constructions: *Any linguistic pattern is recognized as a construction as long as some aspect of its form or function is not strictly predictable from its component parts or from other constructions recognized to exist. In addition, patterns are stored as constructions even if they are fully predictable as long as they occur with sufficient frequency* (Goldberg 2006: 5).

From the perspective of multimodality, this extended definition by Goldberg suggests that there are two forms of evidence for multimodal constructions: (1) a gesture or some non-verbal feature is recurrently used with a given verbal structure and its meaning contribution to the multimodal instantiations is 'not strictly predictable'. This corresponds to the 'strong definition of constructions'; (2) in addition, a pairing of a given gesture (or facial expression etc.) with a given verbal construction could also be stored as a unit if they co-occur with 'sufficient frequency'. As Schoonjans (2014) argues along the lines of Stefanowitsch (2009), Zima (2014a) and many more, both non-predictability and 'sufficient frequency', however, do not provide univocal guidelines for identifying constructions. Especially at the periphery of the phenomenon, i. e. in cases that are not extremely frequent, and for patterns of use that are semantically compositional, but show

idiosyncratic preferences for e. g. certain text genres, sequential positions, or register (Imo 2015), the definition of constructions has been criticized for not being operationalizable (Traugott and Trousdale 2013). This problem of definitional vagueness (for a critical account of both strong and weak definitions of constructions see also Turner) affects traditional, verbally-oriented and cognitive, multimodal C × G alike but it does seem to "come to the fore with greater saliency when considering the possible role of other modalities or modes in the theoretical framework" (Cienki). A number of issues that more or less directly relate to the lack of a univocal definition of the term 'construction' become apparent:

- Is there a frequency of co-occurrence (relative and absolute), which proves the entrenchment of a given construction and a co-verbal kinesic feature?
- Can only a 100 % co-occurrence rate be considered evidence of entrenchment?
- How does this link to the idea of degrees of entrenchment and its gradual nature (Langacker 1987)?
- Some versions of C × G allow for optional elements of verbal constructions. Are kinesic features optional elements of constructions? Is optionality evidence for a lack of entrenchment? If not, how can optional elements be shown to be entrenched?

Papers by Jana Ningelgen & Peter Auer, Alexander Ziem, Thomas Hoffmann, Alan Cienki, Jens Lanwer, Steven Schoonjans, and Elisabeth Zima all provide detailed discussions of the nature of constructions and the different criteria for constructionhood that have been identified in traditional, i. e. verbally-oriented C × Gs against the light of the uncontested multimodality of language use.

The papers by Ningelgen & Auer and Ziem both take a critical stance and align with the 'strong definition of constructions', according to which constructions are structurally and/or semantically unpredictable wholes. They entertain the position that kinesic features must be obligatory, otherwise they cannot be claimed to be entrenched. To test obligatoriness, Ziem proposes to use deletion tests. Only if the suppression of a gesture or other kinesic usage features leads to the uninterpretability of the construction, the kinesic part can be claimed to be an inherent part of that construction. As a consequence, both Ningelgen & Auer and Ziem speculate that the set of truly multimodal constructions (=gesture is obligatory, not optional) is very likely to be very restricted in size.

This contrasts with the approaches that are presented in the papers by Jens Lanwer, Steven Schoonjans and Elisabeth Zima. Their common points of departure are Langacker's definition of entrenchment as a gradual notion and the prototype approach to constructional knowledge. In line with the basic premise of C × G that the constructicon is redundant and constructions are stored with different degrees of schematicity, Jens Lanwer argues that "differences between multi- and unimodal constructions can always be a matter of schematicity". Accordingly, the very same construction could be stored as a very abstract, uni-modal pattern *and* on a more specific level at which information on conventionalized co-speech gesture use is retained. Following up on this, Lanwer considers it unnecessary to decide on whether there is a threshold frequency for gesture occurrence with a construction that points towards its entrenchment (for a discussion see Schoonjans, as well as Schoonjans et al. 2015 and Zima 2014a). Also, Zima argues that it may not be possible to define such a threshold and instead speculates that constructions may be placed on a continuum from constructions which are only infrequently and loosely connected to co-speech gesture use to constructions which are frequently and systematically co-instantiated with a given gesture. Steven Schoonjans takes the argument one step further. He claims that the problems in arguing for entrenchment based on frequency of gesture co-occurrence with a given construction are inherited from mainstream, i. e. verbal-only, C × G and its struggles to operationalize the frequency criterion or the 'weak definition of constructions', respectively. These issues need to be solved in mainstream C × G, before we abandon attempts to explore the multimodality of constructions.

Alan Cienki's paper also picks up on the issue of whether constructions have optional elements that can but do not need to be instantiated in use and suggests to rethink C × G as Utterance Construction Grammar (UC × G). His proposal "starts from the idea of considering the *utterance* as the entry point for characterizing constructions, rather than assuming that gestures should be plugged into existing verbally-based

frameworks, or conversely, assuming that gestures form the higher level of constructional structure, and words fit into them" (his emphasis). Building on Adam Kendon's definition of utterances and Langacker's approach to usage events, Cienki makes a proposal which ironically plays with concepts that are traditionally associated with Universal Grammar: constructions may have a deep as well as a surface structure. A construction's deep structure is conceptualized "as a set of tools that can be drawn upon to express the construction (surface structure)". Whereas, the surface structure is seen as "a metonymic representation of some (if not all) elements of the construction", information about which gestures go with a construction may be part of a construction's deep structure. In this view, every construction hence has a multimodal potential and some aspects of this potential may get activated in a particular usage event and appear at the surface, the level of instantiation, while others may not. This proposal shifts the attention from whether a kinesic feature is entrenched as part of the construction to how central or peripheral it is to the Deep Structure of a given utterance construction. Similar to Lanwer's and Schoonjans' paper, Alan Cienki's programmatic proposal thus attributes prototype theory a central role in cognizing the nature of (utterance) constructions.

Thomas Hoffmann's approach is situated somewhere in between the strong definition of multimodal constructions in terms of obligatoriness of multimodal usage features and proposals that allow for different degrees/levels of multimodality in constructions. He takes both the salience of a gesture and the "frequency with which gesture and verbal FORM co-occur to express the same MEANING" (his emphasis) to have a positive influence on the joint storage in long-term memory. To test the strength of the association between a given verbal construction and a kinesic feature, he proposes to apply collostructional analysis. To be able to use this statistical methodology, however, we need corpora that are tagged for constructions (at all levels of granularity) and all occurring gestures. These corpora are unfortunately not yet available, as also Hoffmann regretfully admits, but if they will be one day, collostructional analysis will certainly put the whole discussion on empirically more solid ground (as Schoonjans also argues). Hoffmann also draws our attention to the fact that "not all constructs simply realize existing stored constructions", a claim that has been made repeatedly in the field of Interactional C × G. To account for the very fact that many but not all form-meaning correspondences in usage are achieved by appealing to stored pre-fabs, he proposes to shift attention to the online, cognitive process through which mental pairings of form and meanings are created through cross-modal associations.

This proposal acknowledges the fact that gestures convey meaning in different ways than speech does. McNeill (1992: 19–23) argues that whereas in speech segmented units are produced in a linear fashion with meaning being constantly updated as discourse unfolds (Langacker 2001), gestures are global and synthetic. They do not gain meaning from separable, meaningful parts but rather meaning is attributed to the parts of gestures only once its global meaning is established. They also usually encode more than one meaning simultaneously and are furthermore claimed to be non-combinatorial, i. e. single gestures do not combine to form larger, more elaborated gestures with more complex meanings (for a critical position see, however, Sweetser 2009). One of the consequences is that the meanings expressed by gestures often do not map onto single verbal expressions, but have larger scope. This is an especially critical issue for Multimodal C × G as constructions are defined as dynamic and flexible, yet *bounded* units.

Accordingly, if a gesture's meaning is situated at another level than the meanings expressed by the verbal construction it co-occurs with, one would probably not treat it as a part of the construction, even if the co-occurrence of the verbal structure and the gesture is recurrent in usage. Another prerequisite for a multimodal construction thus seems to be semantic co-expressiveness (Hoffmann). However, as mentioned earlier, even co-expressive gestures (iconics, metaphorics) usually encode slightly different meanings than those conveyed in speech. On the one hand, it is therefore questionable whether the recurrent use of largely co-expressive gestures with a verbally defined construction points towards the existence of a multimodal construction or not. On the other hand, it may follow that, given the fact that gestures mostly add to the *in situ* meaning of constructions, one should probably be more cautious in calling them 'optional' (Zima, Schoonjans) even if they do not occur in all instantiations.

The prerequisite for answering these questions and all relating issues is data on the (un)systematicity of gesture-construction co-occurrences. Unfortunately, the empirical basis we have so far is still rather weak, as only very few studies provide data for construction-gesture co-occurrences in language use. Yet, C × G is a bottom-up approach and the answers to the questions that are imposed on C × G by the inherent multimodality of language use cannot be found in a top-down manner.

3 Bottom-up means multimodal case-studies first

Several case studies on English, German, and Dutch data are part of the *Linguistics Vanguard* special issue on Multimodal C × G. They contribute to our understanding of how consistently verbal constructions are combined with given gestures in multimodal language use by either focusing on particular verbal constructions (see the papers by Jens Lanwer, Irene Mittelberg, Jana Ningelgen & Peter Auer, Cristóbal Pagán Cánovas & Javier Valenzuela, Mark Turner, and Elisabeth Zima) or a recurrent gesture or bodily conduct with a conventionalized meaning (Bressem & Müller, Jehoul, Brône & Feyaerts, Oakley).

Jens Lanwer focuses on narrow and loose appositions in spoken German, paying special attention to their joint instantiation with head nods. He identifies specific linguistic and para-linguistic characteristics of narrow appositions (e. g. "°hh minIsterpräsident CA:Rstensen kÜndigte AN," 'prime minister Carstensen announced') versus loose appositions (e. g. "[meinen GAST], (.) [pEter SLOterdijk]" 'my guest| Peter Sloterdijk'), and proposes two different constructional schemata for both types. Zooming in on co-speech head nods, Lanwer presents first observations that seem to indicate that these two constructions co-occur with distinct gestural patterns. His data suggest that different kinds of stress-related head nods are associated with some instances of appositions of the narrow type. Lanwer tentatively proposes that these multimodal peculiarities should be "modelled by way of positing a sub-construction of the narrow pattern which would be more specific with regard to both verbal and gestural features".

The paper by Jana Ningelgen and Peter Auer contrasts deictic and non-deictic uses of German *so* (like this/that) with respect to their (non)-dependency on multimodal co-instantiation of manual gestures. Their point of departure is Streeck (2002) and his claim that *so* – irrespective of whether it is used deictically or not – is regularly coupled with an iconic gesture and speaker gaze to his/her gesturing hands. Based on conversational eye tracking data, Ningelgen & Auer refute this very general claim and show that deictic and non-deictic *so* behave very differently from a multimodal point of view. Deicitc *so* is always stressed, accompanied by an iconic/pointing gesture and speakers' as well as recipients' gaze at the speakers' gesturing hands. By contrast, when used as a vagueness/focus marker, *so* is unstressed, iconic gestures and vagueness gestures can be co-instantiated, but they are not mandatory, and neither speaker nor recipient gaze is required. They conclude that only deictic *so* qualifies as a truly multimodal construction whereas for non-deictic *so*, gesture and gaze only function as multimodal metapragmatic cues.

Irene Mittelberg investigates German existential constructions containing the lexicalized chunk *es gibt* (there are). Her aim is to show "how basic manual actions of giving and holding and the corresponding schematic schemes […] motivate multimodal instantiations of existential construction in German discourse". She identifies two gestures from the Palm-up open-hand gesture family (Müller 2004) that co-occur with variants of the *es gibt*-construction: uni-manual PUOH-gestures present (abstract) objects while bimanual gestures hold, enclose or group together objects in-between the two hands. These gestures differ in form and function from full-fledged manual actions of giving. This, Mittelberg argues, may be interpreted as "physical reflections of emergent grammar (Hopper 1998) and *embodied grammaticalization* (Mittelberg and Mortelmans 2013) in gestures that are integrated in multimodally instantiated existential constructions" (her emphasis). Although her study is primarily qualitative in nature and no claims about frequency are made, Mittelberg reasons that "linguistic constructions that recruit basic embodied manual actions and interactions with the physical and social world are particularly likely to be instantiated multimodally and thus also engender emergent multimodal patterns, or clusters, of experience."

This claim is backed up by Elisabeth Zima's study of the [all the way from X PREP Y]-construction. She analyses 199 instantiations of this construction in television data from the Newscape Library of International Television News Broadcasts (Steen and Turner 2013) and shows that 80.4 % of the examples in her data set involve speakers performing a gesture which is co-produced and co-expressive with the verbal construction. However, the frequency of co-speech gesture use significantly varies depending on the use of the construction as either referring to a spatial concept, such as a geographical area, and temporal concepts or list or spectrum of things. Whereas spatial uses are highly prone to co-speech gesture, the latter two are much less so. She further identifies the TV genre that the use of the construction is embedded in, and the preposition in the prepositional slot as factors that have statistically significant influence on the frequency of co-speech gesture use (whereas prosodic stress and repetition of the construction in the immediate discourse context have no such influence). Her study can be seen as an attempt to move beyond mere absolute and relative frequency of co-speech gesture use and to provide a more detailed account of linguistic and contextual usage factors that play a role in whether speakers instantiate a given construction uni- or multimodally.

Cristóbal Pagán Cánovas's & Javier Valenzuela's contribution focuses on gesture use with English temporal demarcative constructions that involve a prepositional slot [from X to Y]. Their main argument is that constructions of this type (e. g. *from beginning to end*) rely on the activation of the timeline as a conventionalized and entrenched template for conceptual integration. This becomes visible in manual gesture use, which they found to be present in up to 80 % of instantiations. Similar to Zima's and Schoonjans's position, they argue that this high frequency is indicative of "the high degree of entrenchment of the gestural information associated with the construction".

Mark Turner's contribution concentrates on a specific type of construction that crucially relies on blending. More specifically, he focuses on multimodal form-meaning pairings for blended classic joint attention (BCJA). He shows that audiovisual media rely on BCJA: Speakers and recipients are not jointly and consciously attending to shared objects as part of their shared ground, but the TV setting and all its agents and peculiarities are blended with our experience of classic joint attention into a scene of BCJA. Turner's paper provides a range of linguistic deictic and non-deictic constructions that prompt for blended classic joint constructions (such as "thank you for joining us"). One particularly interesting non-linguistic example comes from TV news presentations in which an anchor's and a reporter's visual sagittal gazes are parallel as they stand shoulder-to-shoulder, which is blended to mean that the two are looking at each other. As Turner claims, "this multimodal form-meaning pair [is] entirely standard and recognized as unproblematic for TV broadcast news [but] impossible for classic joint attention. It is a canonical form-meaning pair of BCJA in opposition to the classic joint attention form-meaning pair."

The multimodal performance of agents of media discourse is also the topic of Todd Oakley's contribution LINK. He explores two instances of speakers publically engaging in fictive interaction (Pascual 2014) with an absent interlocutor. Oakley argues that fictive interactions "may be considered special types of grammatical constructions, since they are form-meaning pairings that are not only employed in discourse but their representational content is used strategically as a means of structuring whole rhetorical situations". To illustrate his point, Oakley first presents an instance of what he calls "fictive interaction gone wrong" and zooms in on the performance of actor Clint Eastwood at the 2012 Republican Party's National Convention, during which he was engaging in some fictive interaction with President Obama. The failed presentation by Eastwood is contrasted with a similar, yet more convincing scene of fictive interaction between former Governor of South Carolina, Mark Sanford, and Nancy Pelosi. As Oakley argues, the peculiarities of these multimodal performances serve to illustrate the importance of "thick descriptions of individual performances with the aid of new media technologies, combined with a sensitivity to matters such as exigencies of the occasion, the different audiences, and various constraints" as they are "crucial parts of our knowledge of a language's constructicon".

Jana Bressem & Cornelia Müller approach the topic from a slightly different perspective. While most studies that explore the potential multimodality of constructions start from verbal constructions, they depart from a specific recurrent gesture: the Throwing Away gesture. It is "characterized by a particular

kinesic core: a lax flat hand oriented vertically with the palm facing away from the speaker's body flapping downward from the wrist". This form is grounded in our everyday embodied experience of throwing concrete entities away. When the gesture is used with speech, it extends towards metaphorical uses: "The hand now metaphorically throws away topics of talk and qualifies the rejected arguments, ideas and actions as uninteresting and void." Based on a qualitative data study on how this gesture is used in verbal discourse, Bressem & Müller identify a verbo-kinesic pattern called "the negative assessment construction" with the multimodal form [throwing away gesture] + [particles/negation/N/V/ADV]. Drawing on Andrén (2010) and his distinction of item-based and flexible multimodal constructions, they argue that the connection between the gesture and the verbal construction can be more or less flexible, depending on what kind of verbal elements fill the slots in the constructional schema. The paper provides novel ways to think about "multimodal constructions" both by its gesture-first approach and the identification of what might be a network of constructions more or less bound to particular gestures.

Finally, the paper by Annelies Jehoul, Geert Brône & Kurt Feyaerts focuses on the kinesic part of what might be either a uni-modal, gestural, or multimodal, verbo-gestural, "construction of obviousness". They analyze data from a corpus of dyadic interaction between male Flemish speakers, showing that an important role in the gestural marking of obviousness is played by shrugs. Drawing on recent work by Streeck (2009) and Schoonjans (2014), they illustrate that shrugs can consist of a range of gestural components, such as the opening of the hands, shrugging of the shoulders, tilting and shaking of the head and raising of the eyebrows. These different aspects get combined in discourse with different frequencies, with 'complete shrugs' that involve all possible gestural features, being very rare. The very fact that at least in some of their example the epistemic stance of obviousness seems to be solely expressed by (parts of) the shrug without verbal stance markers may point towards the existence of a unimodal, gestural construction

4 Multimodal construction grammar, verbal construction grammar or another model? Where do we go from here?

The special issue on Multimodal Construction Grammar in the *Linguistics Vanguard* is the very first explicitly devoted to Multimodality and Construction Grammar. It brings together different conceptions of what counts as a construction and per extension as a multimodal construction, while staying in the relatively well defined, narrow scope of cognitive versions of C × G. Many questions are raised and discussed, and they certainly outweigh the number of concrete answers. Nevertheless, we believe that the individual contributions of this issue and the links between them greatly add to our understanding of the various issues involved in bringing together multimodal language use and C × G as a theory of the knowledge of language.

References

Andrén, M. 2010. *Children's gestures from 18 to 30 months*. Lund University, Centre for Languages and Literature PhD thesis.

Bergs, A. 2015. *Oh, come on! Problems and perspectives for multimodal construction grammar*. Presentation at Mind, Meaning, Multimodality, University of Navarra, Pamplona.

Blumenthal-Dramé, A. 2012. *Entrenchment in usage-based theories: What corpus data do and do not reveal about the mind*. Berlin: Walter de Gruyter.

Bybee, J. 2006. From usage to grammar. The mind's response to repetition. *Language* 82. 711–733.

Cienki, A. 2012. Usage events of spoken language and the symbolic units we (may) abstract from them. In J. Badio & K. Kosecki (eds.), *Cognitive processes in language*, 149–158. Bern: Peter Lang.

Cienki, A. 2015. Spoken language usage events. *Language & Cognition* 7. 499–514.

Deppermann, A. 2011. Konstruktionsgrammatik und Interaktionale Linguistik: Affinitäten, Komplementaritäten und Diskrepanzen. In A. Lasch & A. Ziem (eds.), *Konstruktionsgrammatik III. Aktuelle Fragen und Lösungsansätze*, 205–238. Tübingen: Stauffenburg.

Deppermann, A. & N. Proske. 2015. Grundeinheiten der Sprache und des Sprechens. In C Dürscheid & J. H. Schneider (eds.), *Handbuch Satz, Äußerung, Schema*, 17–47. Berlin: de Gruyter.

Fillmore, C., P Kay & M. O'Connor. 1988. Regularity and idiomaticity in grammatical constructions: The case of 'let alone.'. *Language* 64. 501–538.

Fried, M. 2015. Construction Grammar. In A. Alexiadou & T. Kiss (eds.), *Syntax – theory and analysis. An international handbook. Handbooks of linguistics and communication science*, 974–1003. Berlin: Mouton de Gruyter.

Goldberg, A. E 1995. Constructions: A construction grammar approach to argument structure. Chicago: University of Chicago Press.

Goldberg, A. E. 2006. *Constructions at work: The nature of generalization in language*. Oxford: Oxford University Press.

Goldin-Meadow, S. 2003. *Hearing gesture: How our hands help us think*. Cambridge, MA: Harvard University Press.

Goldin-Meadow, S. 2013. How our gestures help us learn. In C. Müller, A. Cienki, E. Fricke, S. H. Ladewig, D. McNeill & S. Teßendorf (eds.), *Body–language–communication: An international handbook on multimodality in human interaction*, 792–803. Berlin: De Gruyter Mouton.

Goldin-Meadow, S. & C. Butcher. 2003. Pointing toward two-word speech in young children. In S. Kita (ed.), *Pointing: Where language, culture, and cognition meet*, 85–107. Mahwah, NJ: Earlbaum Associates.

Günthner, S. & W. Imo (eds.). 2006. *Konstruktionen in der Interaktion*. Berlin: de Gruyter.

Hilpert, M. 2014. *Construction grammar and its application to English*. Edinburgh: Edinburgh University Press.

Hopper P. Emergent grammar. 1998. In The new psychology of language: Cognitive and functional approaches to language structure. In: Tomasello M., editors. 155–175. Mahwah, N.J. & London: Lawrence Erlbaum Assoc.

Imo, W. 2007. *Construction grammar und Gesprochene-Sprache-Forschung: Konstruktionen mit zehn matrixsatzfähigen Verben im gesprochenen Deutsch*. Tübingen: Niemeyer.

Imo, W. 2015. Was ist (k)eine Konstruktion? In C. Dürscheid & J. G. Scheider (eds.), *Handbuch, Satz, Äusserung, Schema*, 551–576. Berlin: Mouton de Gruyter.

Kendon A. 2004. Gesture: Visible Gesture as Utterance. Cambridge: Cambridge University Press.

Langacker, R. W. 1987. *Foundations of cognitive grammar: Theoretical prerequisites*, vol. 1. Stanford: Stanford University Press.

Langacker, R. W. 2001. Discourse in cognitive grammar. *Cognitive Linguistics* 12. 143–188.

Langacker, R. W. 2008. *Cognitive grammar: A basic introduction*. Oxford: Oxford University Press.

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

Mittelberg, I. 2014. *Multimodal existential constructions in German and English*. Presentation at GCLA 6, Erlangen, Nurnberg.

Mittelberg, I. & Mortelmans T. 2013. Grammaticalized manual actions of giving and holding: From ditransitive to existential constructions in language and gesture. ICLC 12, Alberta.

Müller, C. 2004. Forms and uses of the Palm Up Open Hand: A case of a gesture family? In C. Müller & R. Posner (eds.), *The semantics and pragmatics of everyday gesture: The Berlin conference*, 233–256. Berlin: Weidler Verlag.

Müller, C. & J. Bressem 2014. *The "Negative-Assessment-Construction": A multimodal pattern*. Presentation at GCLA 6, Erlangen, Nurnberg.

Pascual, E. 2014. *Fictive interaction: The conversation frame in thought, language, and discourse*. Amsterdam: John Benjamins.

Schmid, H.-J. 2007. Entrenchment, salience and basic levels. In D. Geeraerts & H. Cuyckens (eds.), *The Oxford handbook of cognitive linguistics*, 117–138. Oxford: Oxford University Press.

Schoonjans, S. 2014. *Modalpartikeln als multimodale Konstruktionen. Eine korpusbasierte Kookkurrenzanalyse von Modalpartikeln und Gestik im Deutschen*. University of Leuven Unpublished dissertation.

Schoonjans, S., G. Brône & K. Feyaerts. 2015. Multimodalität in der Konstruktionsgrammatik: Eine kritische Betrachtung illustriert anhand einer Gestikanalyse der Partikel *einfach*. In J. Bücker, S. Günthner & W. Imo (eds.), *Konstruktionsgrammatik V. Konstruktionen im Spannungsfeld von sequenziellen Mustern, kommunikativen Gattungen und Textsorten*, 291–308. Tübingen: Stauffenburg.

Steen, F. & M. Turner. 2013. Multimodal construction grammar. In M. Borkent, B. Dancygier & J. Hinnell (eds.), *Language and the creative mind*, 255–274. Stanford, CA: CSLI Publications.

Stefanowitsch, A. 2009. Bedeutung und Gebrauch in der Konstruktionsgrammatik. *Zeitschrift für germanistische Linguistik* 21. 565–592.

Streeck, J. 2002. Grammars, words, and embodied meanings. On the evolution and uses of so and like. *Journal of Communication* 52. 581–596.

Streeck, J. 2009. *Gesturecraft. The manu-facture of meaning*. Amsterdam: Benjamins.

Sweetser, E. 2009. What does it mean to compare language and gesture? Modalities and contrasts. In J. Guo, E. Lieven & N. Budwig (eds.), *Crosslinguistic approaches to the psychology of language: Studies in the tradition of Dan Isaac Slobin*, 357–366. New York: Psychology Press.

Traugott, E. & G. Trousdale. 2013. *Constructionalization and constructional changes*. Oxford: Oxford University Press.

Valenzuela, J. 2015. *Multimodal timelines: Language and gesture*. Presentation at Mind, Meaning, Multimodality, University of Navarra, Pamplona.

Zima, E. 2014a. English multimodal motion constructions. A construction grammar perspective. *Studies van de BKL – Travaux du CBL – Papers of the LSB*, vol. 8. http://uahost.uantwerpen.be/linguist/SBKL/sbkl2013/Zim2013.pdf

Zima, E. 2014b. Gibt es multimodale Konstruktionen? Eine Studie zu [V(motion) in circles] und [all the way from X PREP Y]. *Gesprächsforschung – Online-Zeitschrift zur verbalen Interaktion* 15. 1–48.

Zima, E. In press. Multimodal constructional resemblance. The case of English circular motion constructions. In F. Ruiz De Mendoza, A. Luzondo & P. Pérez-Sobrino (eds.), *Constructing families of constructions*. Amsterdam: John Benjamins.