
This article is under copyright. The publisher should be contacted for permission to re-use or reprint the material in any form. For details, see the journal’s website:

https://benjamins.com/#catalog/journals/gest/main
Abstract
In recent years, consensus has been growing that certain co-speech gestures must be incorporated in grammatical theory. Contrary to traditional beliefs, gestural expression has been found to be systematic in various ways, to exhibit some degree of compositional structure and to be to a certain extent language-specific. Most current work on the gesture-grammar interface is, nonetheless, rather programmatic by character, or lacks an appropriate theoretical framework. This paper shows that Functional Discourse Grammar (FDG) provides a promising model for understanding the grammar-gesture relationship in a more comprehensive and detailed way than has thus far been achieved. It first evaluates how FDG’s perspective on language structure aligns with the ways gestures have been studied. Then it demonstrates that various types of co-speech manual gestures can be (perhaps even exhaustively) classified according the pragmatic and semantic layers recognized in FDG. On the basis of this survey, a first sketch of a multi-channel (speech-gesture) version of FDG’s architecture is presented. The final section puts forward some challenges that arise when taking the multimodal grammar enterprise seriously. Altogether, the content of this paper suggests that a multimodal perspective on grammar can be mutually beneficial for functionally oriented linguists and gesture researchers.

1. Introduction
During spoken interaction, speakers express themselves through various communicative channels, most of which have traditionally remained outside the purview of grammatical theory. Instead, manual and facial aspects of spoken language use have been put forward as paralinguistic or as mere matters of performance (Chomsky 1983). Over the past decades, this bias toward written language has increasingly given way to a more multimodal view on grammar. Along with empirical advances in gesture studies, consensus has been growing that a comprehensive theory of language structure needs to incorporate more than the verbal channel alone.

The aim of this paper is to evaluate the interface between grammar and co-speech gesture from the perspective of Functional Discourse Grammar. The notion gesture is here broadly understood as “actions that have features of manifest deliberate expressiveness” (following Kendon 2004: 15). Although this paper is primarily concerned with manual gestures, many of the arguments presented apply to certain forms of facial behaviors and body posture as well.¹ After reviewing the recent literature on the grammatical potential of co-speech gesture, I show that the various ways in which co-

¹ See Allwood (2003) for a comparative discussion of different types of bodily behaviors from a semiotic perspective.
speech gesture has been studied align well with the main categories and distinctions that FDG postulates. Finally, I sketch a multi-channel (speech-gesture) version of FDG, and discuss some challenges that arise when taking the multimodal grammar enterprise seriously.

2. The grammatical potential of gesture

McNeill (1992: 19) characterizes co-speech gesture as a mode of expression that is “opposite” to speech in certain ways. Whereas speech is segmental, combinatorial and systematic, he argues, gestures are “global-synthetic” and “idiosyncratic” by character, conveying meaning in a “holistic” fashion. This view, which has exerted strong influence on gesture studies, appears at odds with a grammatical approach to gesture, corroborating Chomsky’s (and others’) skepticism on this matter:

Presumably there is a system, a set of principles, that determines the nature of the gestural system and the way in which it interacts with the language system, but whether those principles should be called a grammar is another question. I would think that it is a dubious metaphor, because it leads one to expect commonality of structure, and that is very much an open question. (Chomsky 1983: 41)

In recent decades, nonetheless, evidence has accumulated that gestural modes of expression have more in common with speech than traditionally supposed. In this section I discuss three ways in which gestures are relevant to theories of grammar.

2.1 Systematicity

A necessary precondition for an element of expression to be a candidate subject of grammatical theory, most functional linguists would agree, is for it to have a certain degree of systematicity in the relation between its form and function. Under this criterion, different types of gesture can be ascribed some degree of grammatical potential. Kendon (1988) classifies various types of manual behavior in terms of how conventional or sign-like (in semiotic terms) they are. Of those manual behaviors that occur in concurrence with speech, so-called emblems or quotable gestures are at the most systematic
end of the spectrum. These include gestures such as the thumbs-up, which generally conveys the meaning of positive evaluation, even in the absence of speech, at least in many cultures in the Western world (Efron 1972; Ekman & Friesen 1969; Kendon 1988). The class of speech-linked gestures (e.g., those that complement speech in deictic utterances; McNeill 1992; McNeill 2000) is somewhat more variable in its forms and meanings. Pointing gestures, for instance, can be performed in various ways and can serve a range of different semantic and interactive functions, depending on the verbal and cultural context in which they are deployed (Bavelas 1992; Kita 2003).

Even some forms of spontaneous gesticulation, i.e. those gestures that are produced spontaneously along with speech, are not fully idiosyncratic. Despite the fact that gesticulation is profoundly creative and tailored to the context of the conversation, some patterns and clusters of such gestures have been identified. These have been coined gesture families (Kendon 2004) or recurrent gestures (Müller 2004; Müller et al. 2013b). The cyclic gesture, for instance, whereby the hand makes a repetitive outward rotating movement around the wrist, is linked to a limited set of pragmatic contexts, namely word search, description and request (Ladewig 2011). Based on thorough examination of German multimodal discourse, over fifteen such recurrent patterns in spontaneous gesticulation have been identified and documented (Müller et al., 2013). Notably, recurrent gestures fundamentally differ from emblems in having a flexible form and a meaning that is partly contingent upon the accompanying speech. Cienki (fc.) therefore calls such gestures “semi-symbolic” and represents the relevant part of Kendon’s continuum as follows:

emblems – speech-linked gestures – recurrent gestures – spontaneous gesticulation

more systematic/conventional less systematic/conventional

Figure 1. Different types of gesture ordered according to their degree of conventionality, according to Cienki (after Kendon, 1988).

Experimental research has provided additional evidence that gestures encode semantic and pragmatic meanings in a relatively systematic fashion. Numerous studies have found that the form of
co-speech gestures systematically reflects the discourse structure (Levy & McNeill 1992), linguistic viewpoint (Parrill 2010; Stec 2012), the type of semantic information conveyed (Beattie & Shovelton 1999), and many more such factors (see McNeill 1992; 2000). Of course, one should be cautious to conclude from this body of research that gestural expression is even remotely as crystallized as the verbal component of language. McNeill (1992) was not wrong in stating that speech and gesture are quite different in terms of their degree of discreteness and autonomy as a semiotic channel. In spite of these differences, however, it stands to reason that gestures demonstrate some degree of systematicity on various levels of analysis.

2.2 Language specificity in gesture

A question of particular interest to typology-oriented grammarians is to what extent gestural expression is language-specific. A substantial amount of cross-cultural work has been conducted to address this question (for a review, see Kita 2009). One rather robust finding is that speakers of different languages have distinct repertoires of emblematic gestures (Efron 1972; Ekman & Friesen 1969; Kendon 1992; Kendon et al. 1981; Payrató 1993; Sparhawk 1978). In Dutch, for example, indicating that the food that is one is eating is tasty is done by rapidly waving a hand next to the cheek, but in Italian, this is done by twisting the index finger back and forth into the cheek. Comparative inquiry suggests that the way people use their body for making deictic reference is not stable across cultures either. Many different formal variants of index finger points are known (Kendon & Versante 2003; Wilkins 2003), as well as cases of pointing by means of the lips (Enfield 2001; Sherzer 1973) and the nose (Cooperrider & Núñez 2012).

Intercultural differences in gesturing have furthermore been observed on a more abstract level, for instance in the way people use their hands while presenting spatial descriptions. Whereas in most cultures people use their own body as a reference point when using gestures to represent spatial scenes, records exist of communities that deploy an absolute frame of reference when doing so (Haviland 1993; Levinson 2003). These findings, notably, are only relevant to the current discussion to the extent that they reflect communicative conventions rather than conventions in imagery. It is important to be aware that regularities in gestural expression do not always reflect linguistic patterns,
but could instead emerge from commonalities in the way people conceptualize their environment (Okrent 2002; Parrill, 2010).

Nonetheless, the existence of culture-specific, communicative patterns in gestural expression is evident, especially when it comes to emblems and indexical gestures. At least to some extent then, co-verbal expression involves learned and partly arbitrary knowledge, manifested in codified symbols as well as more schematic expressive patterns.

2.3 Gestures in intersection with the syntax of speech

Both in terms of the cognitive processes underlying their expression (McNeill & Duncan 2000) as well as in terms of their role in communication (Enfield 2004; 2009; Kendon 2004; Streeck 2009), speech and gesture are best thought of as strongly intertwined rather than as autonomous mechanisms. As Kendon (2004: 127) puts it, “[i]n creating an utterance that uses both modes of expression, the speaker creates an ensemble in which gesture and speech are employed together as partners in a single rhetorical enterprise” (emphasis in original). Accordingly, gestural expression should not be treated as a mere addendum to the grammar of the verbal channel, but as an integral part of linguistic structure.

One manifestation of this interaction is gesture’s potential to integrate with the syntax of speech. Speech and gesture may, for instance, combine in a sequential fashion (Enfield 2009; Ladewig 2012; Slama-Cazacu 1976; Streeck 1993), whereby gestures substitute verbal elements and take over their syntactic function. Examining the occurrence of gestures in cases of interrupted speech, Ladewig (2012) found that gestures most often occupy noun and verb slots in utterance-final position. Exactly what type of grammatical element a gesture substitutes cannot always be conclusively determined, however. The gesture in (1) (adopted from McNeill 2005: 5) can for instance be interpreted as replacing either a non-finite ing-clause (e.g. ‘flying out the window’) or a prepositional phrase (e.g., ‘into the garden’ or ‘like this’).

(1) Sylvester went… [gesture of an object flying out laterally]
Another way gestures integrate with speech is by providing semantic content while being produced in concomitance with the verbal channel. Fricke (2008; 2009; 2012) reports that gestures, when they coincide with speech, often take the role of an attribute to a nominal phrase. In such cases, a gesture can either refine verbally presented information (2) or present additional information (3).

(2) A small wooden sculpture [opposing the index finger and thumb to give a more precise qualification of the size of the sculpture, thus elaborating the verbally encoded property *small*]

(3) A small wooden sculpture [tracing a line in the air to show the contours of the shape of the sculpture in the air (a property which is not expressed verbally)]

What follows from this work is that manual gestures, even those that are not themselves conventionalized, can intersect with the grammar of speech. In addition, recent work has shown that gestures produced in concurrence with speech reflect higher-level grammatical functions. Verbal expressions of grammatical aspect (Bressem 2012; Duncan 2002), negation (Harrison 2009) and modality (Schoonjans 2014) tend to go along with specific gestural patterns in certain languages. In some cases, the performance of a gesture even seems obligatory, in the sense that omitting it would render the utterance infelicitous. Utterances like *look over there*, and other forms of spatial deixis, are only felicitous when accompanied by a pointing gesture of some form (Kita 2003; Streeck 2002). These findings call for an approach to grammar that accommodates, or at least acknowledges, the linguistic functions of co-speech gestures.

3. A Functional Discourse Grammar perspective

The points of convergence between grammar and gesture research outlined above have motivated some linguists to entertain a *multimodal* conception of grammar. Protagonists of Systemic Functional Grammar have been pioneers in this line of investigation, first concentrating on text-image combinations (e.g. Kress 2000; Kress & van Leeuwen 1996), and later taking into consideration the relation between spoken language and gesture (Muntigl 2004). Other accounts have adopted the

The last-mentioned Functional Discourse Grammar (henceforth FDG; Hengeveld & Mackenzie 2008) combines a distinct number of features which make it an interesting candidate for further pursuing this avenue of inquiry. First of all, FDG is designed to include semantic and pragmatic factors simultaneously in grammatical analysis. Since gestures typically carry out representational and interactive functions in tandem (Müller 2013), this is an important prerequisite for an adequate functional model of multimodal language structure. Second, FDG allows for grammatical analyses to be carried out in a comprehensive manner. In contrast to constructionist approaches to grammar, FDG’s basic unit of analysis is a full interactional Move (i.e., a unit of interaction, e.g. a conversational turn) rather than an individual grammatical element or construction. This enables a close scrutiny of the function of a gesture in relation to a detailed functional-structural analysis of the verbal channel. Finally, FDG provides a very rigorous analytical formalism that enforces a high degree of explicitness in its analyses.

Nonetheless, the application of FDG to multimodal data is still in its infancy. Connolly (2010), in the first paper to discuss FDG’s multimodal potential, concludes that FDG is in principle compatible with a multimodal conception of discourse, provided that it is understood as a branch of semiotic theory. Like many other assessments of the gesture-grammar interface, however, Connolly’s analysis has remained rather programmatic and discusses only a few ways in which gesture (and other non-verbal modes of communication) can be relevant to the study of grammar.

With the aim of achieving a more detailed view of gesture’s grammatical functions, the remainder of the present paper evaluates a range of connections between gesture research and language structure as conceived by FDG. The following subsections show that the various ways in which co-speech gesture has been studied align well with the main layers of interpersonal and semantic representation postulated in FDG. These are summarized in Table 1.
Table 1. A summary of the most important semantic and pragmatic layers recognized in FDG

<table>
<thead>
<tr>
<th><strong>Level of representation</strong></th>
<th><strong>Definition</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Interpersonal level</strong></td>
<td></td>
</tr>
<tr>
<td>Move</td>
<td>“an autonomous contribution to the interaction” (Hengeveld &amp; Mackenzie 2008: 50)</td>
</tr>
<tr>
<td>Discourse Act</td>
<td>“the smallest identifiable units of communicative behavior” (Hengeveld &amp; Mackenzie 2008: 60; after Kroon 2005)</td>
</tr>
<tr>
<td>Illocution</td>
<td>“the conventionalized interpersonal use [of a Discourse Act] in achieving a communicative intention” (Hengeveld &amp; Mackenzie 2008: 68)</td>
</tr>
<tr>
<td>Communicated Content</td>
<td>“the totality of what the speaker wishes to evoke in his/her communication with the addressee” (i.e., the actional dimension of expressing a Propositional Content; Hengeveld &amp; Mackenzie 2006: 376)</td>
</tr>
<tr>
<td>Ascriptive Subact</td>
<td>“an attempt by the Speaker to evoke a property” (i.e., the actional dimension of expressing a Lexical Property; Hengeveld &amp; Mackenzie 2006: 376)</td>
</tr>
<tr>
<td>Referential Subact</td>
<td>“an attempt by the Speaker to evoke a referent” (i.e., the actional dimension of referring to some entity; Hengeveld &amp; Mackenzie 2006: 376)</td>
</tr>
<tr>
<td><strong>Representational level</strong></td>
<td></td>
</tr>
<tr>
<td>Propositional Content</td>
<td>“mental constructs, such as pieces of knowledge, beliefs, and hopes” (Hengeveld &amp; Mackenzie 2006: 377)</td>
</tr>
<tr>
<td>Episode</td>
<td>“sets of States-of-Affairs that are thematically coherent” (Hengeveld &amp; Mackenzie 2006: 378)</td>
</tr>
<tr>
<td>State of Affairs</td>
<td>“entities that can be located in relative time and can be evaluated in terms of their reality status” (Hengeveld &amp; Mackenzie 2008: 166)</td>
</tr>
<tr>
<td>Configurational</td>
<td>The internal organization of a State of Affairs.</td>
</tr>
</tbody>
</table>
**Property**  
(Hengeveld & Mackenzie 2006: 379-380)

<table>
<thead>
<tr>
<th><strong>Lexical Property</strong></th>
<th>A (lexically expressed) characteristic of a semantic unit (Hengeveld &amp; Mackenzie 2008: 215-236)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Individual</strong></td>
<td>“a concrete, tangible entity” (Hengeveld &amp; Mackenzie 2008: 236)</td>
</tr>
</tbody>
</table>

All of these layers have a hierarchical organization themselves. Each layer is restricted by a head and optionally one or more *modifiers* (lexical by nature), and further specified by *operators* (grammatical by nature). FDG furthermore distinguishes a number of *functions*, which specify relations between units at the same layer. Those functions that may hold within and between Communicated Contents are most relevant for the present paper. These include Focus-Background (pertaining to the degree to which the provided information is new or accessible), Topic-Comment (pertaining to the communicative point of departure for the utterance) and Contrast-Overlap (pertaining to the intention of the speaker to draw attention to some point of discrepancy between elements of the Communicated Content(s) expressed; Hengeveld & Mackenzie 2008: 89-100).

In the following, I attempt to show that the layers of pragmatic and semantic organization summarized above are all known to receive gestural expression under certain circumstances. In the light of this finding, this section concludes with a proposal for a multimodal extension of FDG.

### 3.1 Gesture at the Interpersonal Level

Gestures contribute in various ways to the interpersonal dimension of language use, as follows from experimental (e.g., Bavelas et al. 1992), semiotic-communicative (e.g., Kendon, 2004) as well as praxeological research (Streeck 2009). The comprehensive architecture of FDG makes it possible to bring together these resources into a single coherent framework. Here I outline how some of the most well-known interpersonal-pragmatic functions of gestures can be interpreted according to FDG’s layered hierarchy of pragmatic organization.

An interactional Move can be realized by a gesture alone, for instance when waving one’s hands to greet someone or nodding to approve something. When a Move is expressed by speech, gestures are oftentimes deployed to designate its beginning or end (along with verbal markers).
Streeck (2009) for instance describes how speakers mark the end of a conversational turn – which may encompass one or more Moves – by spreading out their arms toward the addressee, as if literally handing the turn over to the interlocutor.

Gestures with a so-called parsing function are relevant with respect to Discourse Acts. These are gestures that “contribute to the marking […] of the structure of spoken discourse” (Kendon 2004: 225). When summing up a list of points or arguments, for instance, speakers often keep track of the overall structure of the discourse (the respective organization of Discourse Acts, or clusters of them) by holding one hand up and stretching one finger at a time in rough temporal correspondence with the points or arguments presented (Streeck 2009; Sweetser 1998).

Several examples are known of conventionalized gestures that contribute to the Illocution layer (Kendon 1995). A well-documented gestural complement of interrogatives, for example, is the purse hand in Southern Italian discourse, whereby “all the digits of the hand are held fully extended, but they are drawn together so that they are in contact with one another at their tips” (Kendon 1995: 249). Kendon moreover elaborates on the joined hands gesture, which he says to mark the discourse “as an appeal to others to take a certain line of action” (Kendon 1995: 261). This gesture can be analyzed as operating on imperatives, transforming what would otherwise come across as an order into a request or plea. Thus, it has a mitigating role.

In terms of the Communicated Content layer, what is relevant is the close connection between gesture and information structure. Some have claimed that gestures are a direct expression of high communicative dynamism, appealing to Firbas’ (1964) definition as the extent to which an utterance “pushes the communication forward” at some point (Levy & McNeill 1992; McNeill 1992: 207). Accordingly, some gestures may best be analyzable as manifestations of Focus or Comment functions operating on Communicated Contents. Drawing once again on analyses of Southern Italian discourse, Kendon (1995) furthermore characterizes the finger bunch gesture as a marker of topic-comment oppositions in the accompanied speech. The fingers here are first held together and subsequently opened, in close coincidence with the discourse segments that serve topic and comment functions, respectively. The Contrast function, furthermore, may receive rather specific gestural coding in the
form of a weighing gesture, where two Communicated Contents are opposed by holding the two hands in front of the body with the palms faced up, shaped like cups (McNeill & Levy 1982).²

Gestures (or aspects of them) pertinent to the layer of the Referential Subact include those intended to make a referent identifiable to the addressee. In the case someone points while asking *Could I have this one?* (e.g. in a store), the pointing gesture can be analyzed as expressing a Referential Subact (with an identifiability operator), which is co-indexical with the Referential Subact expressed in speech. Gestural correlates of existential constructions can be relevant at this layer as well. Mittelberg & Mortelmans (2013) found that constructions such as *there is* in English and *es gibt* in German are often accompanied by a *palm-up-open-hand* gesture. This recurrent gesture is known to typically “present an abstract discursive object as an inspectable […] and it invites participants to take on a shared perspective on this object” (Müller 2004: 233). The performance of a palm-up-open-hand gesture in combination with an existential construction, according to this characterization, can be hypothesized to co-express (or further refine), the focal status of the Referential Subact.

Gestural components of Ascriptive Subacts, finally, subsume those that express the speaker’s subjective commitment to an ascribed property. Certain types of shoulder shrugs and hand gestures are known to be attuned to this purpose (Debras & Cienki 2012). The *vague gesture* (Müller *et al.* 2013a), whereby the lax flat hand is repeatedly turned relative to the wrist, can be seen as a manual equivalent of approximation markers (e.g. -like, -y and -ish suffixes in English; Hengeveld & Keizer 2011). When indicating the size of an object by holding the hands at a certain distance from each other, for example, holding the hands firm and tensed hints at a fair amount of self-reliance, whereas holding the hands lax and twisting them around the wrist conveys a lesser degree of confidence that the ascribed property is accurate. In these cases, the position of the hands can be compared to a (Lexical) Property at the Representational Level, but the degree of tension and stability of the hands relate to the Interpersonal Level at the layer of the Ascriptive Subact.

### 3.2 Gesture at the Representational Level

² Note that the expression *on the one hand... on the other hand...* plausibly has its roots here.
A substantial body of gesture research is concerned with the semantic functions of gesture, i.e. those that relate to the Representational Level in FDG. Despite the fact that studies on gestural semantics typically draw on ad-hoc categories (e.g. Beattie & Shovelton 1999) or semantic constructs from cognitive linguistics (Cienki 2005), it is possible to classify gestural patterns in terms of their relation to FDG’s representational layers.

Head nods in response to yes-no questions, to begin with, may be analyzed as Propositional Contents. Akin to the words yes and no, head nods can substitute entire Propositional Contents, assigning a truth value to some proposition. Furthermore, some manual gestures relate to the speaker’s degree of commitment to the proposition that is expressed verbally. The vague gesture, mentioned before, has been observed in some cases to “mark […] ideas as uncertain or indeterminate” (Müller et al. 2013b: 720).

Recent research has also hinted at a gestural correlate of Episodes. According to Müller et al. (2013a), the performance of multiple gestures in close spatial or temporal proximity engenders the mental construction of coherent semantic scenarios. When speakers gesture while describing a series of temporally or spatially connected events, they typically do not return their hands to a rest position in between successive gestures. In line with the definition of Episodes in FDG, the physical continuation of kinesic expression in such cases indicates that different states of affairs form a temporally or spatially coherent whole, such as in (4).

(4) I first drove east and then traveled back up north [tracing a line in the air, first horizontal, then vertical, without returning the hands to rest position in between].

Gestures that relate to States of Affairs include those that convey information with respect to their “properties of occurrence” (Hengeveld & Mackenzie 2008: 171). The relative temporal occurrence of two States of Affairs, for instance, can be reflected by the temporal ordering of the two gesture strokes (as in (4)). Gestures that co-express negation, typically characterized by the horizontal outward movement of one hand with one or multiple fingers stretched (Harrison 2009; Kendon 2004), may pertain to this layer as well.
Concerning the layer of the Configurational Property, those aspects of gestures play a role that have to do with the internal temporal organization of States of Affairs. Most relevant here is research on gestural correlates of grammatical aspect. Various studies have reported that speakers (of English) make longer and more complex gestures when producing utterances with progressive-marked aspect than with perfective aspect (Duncan 2002; McNeill 2003; Parrill et al. 2013).

Lexical Properties and Individuals, finally, can be represented gesturally in many ways. Müller (1998) distinguishes a number of strategies people use for accomplishing iconic reference to an object or person, e.g. by molding its shape, tracing its contours or embodying the object so that the hand becomes the object itself. By virtue of this iconic potential, gestures are especially adept at representing information related to shape (3), size (5) and motion (6). Interestingly, Individuals and their corresponding Properties are often encoded in one and the same gesture. In (6), for instance, the hand represents an Individual (the rock) by embodying it, but at the same time (co-)ascribes two Properties to it (shape and spatial trajectory).

(5) I just saw a huge rat! [holding the hands apart at a certain distance, palms facing each other]

(6) The rock bounced down the cliffs [moving a fisted hand through the air to trace the rock’s trajectory]

3.3 Gesture and Encoding
As yet, the structural mechanisms underlying gestural expression are not understood in sufficient detail to make an explicit and detailed proposal on what a gestural Encoding operation entails (i.e., the part of the grammar model that transforms semantic and pragmatic representations into a readily expressible form). Given the intrinsic differences in the communicative channels that speech and gesture exploit, however, it is clear that an adequate multimodal grammar requires multiple, distinct levels of form.

Despite the absence of a single, unitary formalism for describing gesture form, structure-oriented approaches have alluded to the existence of gestural primitives. First of all, contrary to
McNeill’s (1992) holistic view, recent evidence suggests that gestures do have some degree of segmental structure. Researchers focusing on relatively conventionalized gestures (Calbris 1990: ch3; 2008; Kendon 2004; Webb 1996) have advocated the view that gestural expression involves a number of organic kinesic patterns and handshapes (e.g. curves, loops, lines and planes), which are consistently linked to certain semantic contexts. Provided that some formal facets of these gestures can be recombined to create novel meaningful patterns, Calbris (1990) and others have suggested that gestural expression is at least partly compositional by nature (although the kind of compositionality involved here is not of the same linear-segmental nature as verbal syntax).

On related grounds, some form-based approaches to gesture have adopted the four phonological parameters from sign language studies: handshape, orientation, location and movement (cf. Stokoe 1960/2005). Gesture notation systems based on these elementary form features have proved helpful in identifying recurrent patterns in spontaneous gesticulation (Müller et al. 2013a). The observation that spontaneous gestures overlap in their formational make-up has moreover given rise to the suggestion that gesture form is governed by a “rudimentary morphology” (Müller 2004: 3) or an “emerging morphosemantics” (Kendon 2004: 224).

A different line of research has focused on the articulator-referent mappings that underlie iconic representations in spontaneous gesticulation. Based on systematic analyses of the forms of referential gestures in relation to the spatial characteristics of their referents, it has been suggested that creative iconic gestures are not as unconstrained as previously supposed (Bergmann & Kopp 2010; Hassemer et al. 2011; Poggi 2008). The formal make-up of iconic gestures has been characterized in terms of generative (Poggi 2008) as well as probabilistic grammatical formalisms (Sadeghipour & Kopp 2014).

In terms of their temporal-sequential organization, it has furthermore been argued that gestures follow a kind of hierarchical constituency structure (Fricke 2012; 2013). According to Fricke, the way gesture phases (e.g. the preparation of the gesture, the most effortful part or stroke phase, and the retraction of the hand into a rest position; Kendon 1972) combine into bigger units can be captured by a limited number of “syntactic rules”, that are hierarchical and recursive (Fricke 2013: 743-744).
Whether the notion of syntax can be extended so seamlessly to the domain of co-verbal expression, however, remains a topic of ample controversy.

A final notable finding is that speech and gesture encoding processes are in close interaction with each other. A series of studies by Özyürek and colleagues (Kita & Özyürek 2003; Özyürek et al. 2005) showed that simultaneous speech-gesture combinations are sensitive to the typological characteristics of the language in which they occur. Speakers of verb-framed languages (Talmy 1985), which express spatial information on manner and path in separate clauses, tend to break up manner and path into two consecutive gestures. Conversely, speakers of satellite-framed languages, which express manner and path in a single clause, were found to conflate manner and path in one and the same gesture.

Altogether, at least four types of primitives can be hypothetically assigned to a gestural encoding module: (1) lexeme-like gestural patterns such as emblems, (2) basic phonological parameters (handshape, orientation, location and movement) that can be combined productively, (3) morpheme-like patterns such as planes, curves and lines, and (4) patterns underlying iconic mappings. Yet, various questions and points of dispute remain with respect to the findings presented in this section. Despite the existence of some degree of compositionality, it is irrefutable that gestural meaning is strongly constrained in a top-down fashion, i.e. by verbal and situational context. This to some extent sets gestural patterns apart from the verbal primitives postulated in FDG, such as lexemes and syntactic templates, which have stable grammatical and/or semantic properties that are relatively independent of the context in which they are attested.

3.4 Initial observations

The explorations outlined thus far yield some preliminary observations. First of all, there appears to be a difference in the types of gesture that relate to interpersonal layers on the one hand and representational layers on the other. Those gestures carrying out interpersonal functions are almost without exception conventionalized to some degree – they are either emblematic or classifiable as a recurrent gesture. Lower layers at the Representational Level (e.g. Individuals or Properties), by contrast, more often involve creative forms of gesticulation. A plausible explanation for this
divergence is that the hands have a natural potential for creating iconic mappings with respect to semantic aspects of conceptualization, but much less so with respect to interpersonal aspects. That is, the hands can perform referential and ascriptive functions without drawing on a conventionalized code, but instead by exploiting iconic relations between their physical form and the spatial features of the referent. This applies much less to interpersonal functions, which do generally not have inherent spatial or motoric characteristics that allow for creating iconic mappings. The pressure for gestures with interpersonal functions to become conventionalized may therefore be stronger than for those with a semantic function.

Second, gestures’ (grammatical) functions are often dependent on what is conveyed verbally (cf. Feyereisen et al. 1988). Cases where gestures serve as the head of a formulation layer are rare, with the exception of those embodying full Moves or Propositional Contents. In the examples discussed, gestures predominantly provide further qualifications of (or operations on) verbally encoded units.

3.5 A first sketch of a multi-channel FDG

An appropriate model of multimodal grammar, as follows from the literature reviewed thus far, entails separate, but mutually interactive formulation and encoding operations for gesture and speech. The architecture of such a model can be inspired by psycholinguistic models of speech-gesture co-production (e.g. De Ruiter 2000; Kita & Özyürek 2003), which have extended Levett’s (1989) model of language production (which has been a source of inspiration for FDG’s architecture) with a parallel component for gesture. Along the same lines, Figure 2 presents a possible architecture of a multimodal extension of the FDG model, following Connolly’s (2010) suggestion to treat the different modalities as constituting separate, but mutually interactive semiotic systems. Additions relative to the current version of the model are marked with dashed lines.

Some pragmatic gestures, such as the ‘weighing gesture’ described above, are iconic by establishing a metaphorical mapping to a concrete spatial domain. However, such gestures are still conventionalized to some degree; they do not rely on iconicity in the same way as gestures with a referential or ascriptive function.
Figure 2. A candidate architecture of a multimodal extension of FDG

The Conceptual Component remains in essence unchanged relative to current FDG. It contains the totality of the interactional and ideational material that is to be translated into ostensive signs (manifested through the auditory as well as the visual-manual modality). The Grammatical Component here can be understood in a broad, inclusive sense, namely as comprising the totality of socially shared knowledge for enabling contextualized meaning making (through whatever modality). Note that the postulation of a single, multimodal grammar is not meant to imply that the operations underlying speech and gesture generation are of the same kind. Instead, as discussed above, each
modality has its own operations and draws on a separate set of primitives, with distinct structural features and levels of schematicity.

The initial, rough division of communicative labor over verbal and manual modalities is determined during Macro-Planning. The reason for including this operation in the Grammar is the ostensible existence of *multimodal constructions* in certain languages, i.e., fixed speech-gesture combinations that form a single expressive unit (Sandoval 2014; Schoonjans, 2014). 4 Further distribution of labor over speech and gesture can be assumed to be established on the fly during Formulation, as a result of mutual interactions between the Speech Formulation and Gesture Formulation operations (cf. McNeill & Duncan 2000). This distribution is in part determined by the format of the conceptual content activated: conceptual material of a *spatial-motoric* nature is a more natural candidate for being expressed manually, whereas more abstract types of content are most readily encoded by verbal means (cf. Kita 2000).

The Speech Encoding and Gesture Encoding operations further elaborate the output of these processes into a form that is encodable in speech and gesture, respectively. Possible primitives that this operation draws on have been discussed in section 3.3. In keeping with the finding that gestural encoding of information flexibly adapts to the syntax of the verbal channel, a tight interaction can be assumed between these Encoding operations. Additional configurational templates may be required here to explicate under which conditions the sequential combination of speech and gesture is acceptable (e.g., why (7) is more felicitous than (8)). Further elaboration on what these templates look like remains outside the scope of this paper.

(7) Then the car went [tracing the trajectory of the car in the air]

---

4 Another possible example of a Macro-Planning operation is the deliberate creation of irony by presenting mismatched semantic information in speech and gesture (e.g., saying “big” but indicating a small size gesturally). Because in such cases the interpersonal value of the utterance (the irony) emerges from the interplay between the two channels, it must be assigned to an operation that precedes modality-specific formulation. According to FDG’s principle of formal encoding, however, this type of phenomenon is only of interest to the extent that it receives systematic formal expression.
(8) *He [gesture depicting the action of walking] to the shop

The Contextual Component, finally, subsumes all elements of context that have a systematic impact on verbal and/or gestural form. This includes such factors as the (real or fictive) presence of a referent in the immediate environment, the physical position of the interlocutor, the assumed common ground and the gesture space previously used. Some of these factors are specifically relevant to gesture, whereas others may influence speech and gesture expression in similar ways.

It should be kept in mind that Figure 2 is not to be conceived of as a processing model in and by itself. Although inspired by psycholinguistic models of speech-gesture production, it is meant as a framework for multimodal grammatical analysis. The strong resemblance to a production model is in accord with FDG’s assumption that “a model of grammar will be more effective the more its organization resembles language processing in the individual” (Hengeveld & Mackenzie, 2008: 2-3). In keeping with this tenet and the empirical evidence synthesized above, a multimodal FDG needs to reflect at least the following three empirical facts: (1) speech and gesture are components of one and the same communicative process; (2) speech and gesture are mutually interactive during formulation and encoding; (3) yet, speech and gesture involve partially independent primitives and operations. Although the architecture proposed in Figure 2 is certainly not the only possible model one can think of, its explicit incorporation of all three of these empirical facts renders it fruitful for the current purposes.

3.6 Example analyses

To give an impression of how multimodal FDG can be deployed in practice, this section presents two example analyses of composite, spoken-gestured utterances. Both concern excerpts taken from the Bielefeld Speech and Gesture Alignment corpus (SaGA; Lücking et al. 2010), which contains video recordings of people engaged in a route-description task. The parts of the utterances analyzed are marked in boldface. In interest of space, only the following variables are considered: Communicated Contents (C), Referential Subacts (R), Ascriptive Subacts (T), Individuals (x) and
properties (f) (see Table 1 for a description of these variables). The dummy variable $\alpha$ is used as a placeholder for any variable on the Representational Level. Threefold analyses are given on the Interpersonal Level as well as the Representational Level: one for each of the modalities separately, and one for the utterance as a whole. This approach mirrors the assumption that although speech and gesture convey semantic and pragmatic material by themselves, they ultimately represent a single, unitary utterance. Subscript indices mark distinct variables and indicate correspondences between variables at the different levels (e.g., the Referential Subact $R_i$ corresponds to the Individual $x_i$ on the Representational Level). In the combined representations, double subscripts (e.g., $x_{ij}$) indicate co-indexicality between variables on different tiers (e.g., $x_i$ in speech and $\alpha_j$ in gesture). This notation illuminates to what extent speech and gesture present overlapping semantic content. In addition, correspondences between variables are indicated by the horizontal alignment between the symbols on the different tiers (corresponding variables are printed below and above on another).

Both examples (9) and (10) are excerpts of a discourse in which the speaker describes a route through a town to an addressee who has not seen this town before.

(9) \begin{tabular}{lllll}
\textit{in} & \textit{der} & \textit{Mitte} & \textit{ist} & \textit{noch} \\
\textit{in} & \text{ART.DAT.F.SG} & \text{middle} & \text{be.PRS.3SG} & \text{yet/still} \\
\textit{blau-es} & \text{Eingangstor} \\
\text{blue-NOM.N.SG} & \text{entrance.gate} \\
\end{tabular}

‘In the middle there is another \textbf{blue entrance gate} like that.’

\footnote{Also, in order to make the formalism more accessible, closing variables have been omitted}
**Figure 3.** A gesture that depicts the shape of an entrance gate

While saying *blaues Eingangstor* ‘blue entrance gate’ the speaker moves his index fingers first outwards and then downwards, thus tracing the rectangular shape of the entrance gate in the air (Figure 3). This multimodal utterance segment can be analyzed as follows:

### IL

- **speech**
  - (Ci: (Ri) (Ti) (Tj))
- **gesture**
  - (Cj: (Rj) (Tk))
- **combined**
  - (Ck: (Ri) (Ti) (Tj) (Rj) (Tk))

### RL

- **speech**
  - (xi: ((fi: Eingangstor): (fj: blau)))
- **gesture**
  - (αj: (fk: gestured shape))
- **combined**

The spoken part of the utterance segment performs a Referential Subact and two Ascriptive Subacts. These Ascriptive Subacts correspond to the Properties ‘Eingangstor’ *entrance gate* and ‘blau’ *blue* on the Representational Level. The gestural component of the utterance, in addition, performs a Referential Subact and an Ascriptive Subact. The tracing movement of the hands evokes a semantic entity of some kind (presumably an Individual, but possibly a higher-level unit; hence the placeholder α) and ascribes a spatial Property to it (a square shape). Because the verbally and gesturally expressed Referential Subacts are clearly co-indexial, the combined semantic representation contains only a single Individual (xij). In contrast, the Property corresponding to the gesturally expressed Ascriptive Subact is **complementary** with the verbal tier: akin to the word ‘blaues’, the gesture modifies the property ‘Eingangstor’. Thus, according to the analysis given in (9), the gesture has nominal as well as adjectival characteristics. The example shown below in (10) illustrates another combination of gestural and verbal descriptions.
‘You have to **kinda** go between through those **ten trees**’

**Figure 4.** A gesture that designates the quantity of ten and expresses uncertainty

Roughly in concurrence with the articulation of *zehn Bäumen quasi* ‘ten trees kinda’, the speaker raises her both hands with the palms faced towards the addressee and all fingers extended vertically. Thus, her hands co-express the quantity of the trees she talks about (ten). Meanwhile, she repeatedly rotates her hands sideways around the wrist, ostensibly signaling uncertainty as to whether the information she provides is accurate. This uncertainty is expressed in speech as well, namely by the hedging particle ‘quasi’ *kinda*. An FDG analysis of this utterance segment is given below. The parts of the utterance that are not directly relevant to the speech-gesture interaction are omitted.

\[
\begin{align*}
\text{IL}_{\text{speech}} & \quad (\text{approx C}_i: \text{([…])} \quad (\text{R}_i) \quad (\text{T}_i)) \\
\text{IL}_{\text{gesture}} & \quad (\text{approx C}_j: \quad (\text{R}_j)) \\
\text{IL}_{\text{combined}} & \quad (\text{approx C}_k: \text{([…])} \quad (\text{R}_i) \quad (\text{T}_i) \quad (\text{R}_j)) \\
\text{RL}_{\text{speech}} & \quad \text{([…])} \quad (10\alpha_i: \quad (f; \text{Baum})) \\
\text{RL}_{\text{gesture}} & \quad (10\alpha_j)
\end{align*}
\]
The Communicated Content $C_i$ on the spoken tier represents the action of informing the addressee ‘that she has to go through those ten trees’). The speaker’s low commitment to the accuracy of this information, signaled by the hedging particle *quasi* ‘kinda’, is represented by the approximation operator $(\text{approx})$ on the Communicated Content (cf. Hengeveld & Keizer, 2011). The gestural component of the utterance co-expresses this uncertainty: the rotating of the hand palm around the wrist shows that the speaker lacks full commitment to what she says.

In addition, the gesture contributes rather specific semantic information: the vertical raising of all fingers on both hands evokes an entity (of any order) and designates its quantity (ten). The combined analysis on the Representational Level elucidates how this information relates to the semantics of verbal channel. The gesturally evoked entity is ostensibly co-indexical with the Individual that is semantically restricted by the Property $Baum$ ‘tree’ and the quantificational operator $zehn$ ‘ten’ (i.e., the same set of trees is being referred to). Because of this semantic correspondence, the combined Representational Level contains just a single Individual ($x_{ij}$) and a single quantificational operator.

In sum, the gesture interacts with speech on two different levels of representation. It expresses approximation with respect to the Communicated Content – an operation that is relevant at the Interpersonal Level – and it performs a low-level semantic operation: the quantification of an Individual. It is noteworthy that both of these functions of the gesture are also expressed in speech: the former by the hedging particles *quasi* ‘kinda’, and the latter by $zehn$ ‘ten’. Thus, in this example, the gesture does not contribute additional semantic content relative to the speech, like was the case in (9).

These analyses are rather preliminary and many questions and challenges remain. One issue that needs more attention is how the modifier-operator distinction is to be (re)interpreted in the
gestural realm. Further and more general challenges are discussed in Section 4. Nonetheless, the examples in their current form demonstrate some of the strengths of FDG as a tool for multimodal grammatical inquiry. By enforcing a view on gestures as simultaneously actional and representational, FDG enables a detailed understanding of the structural relationship between the verbal and gestural channels of expression.

4. Remaining challenges

Despite the optimism expressed so far, a number of challenges and potential problems need to be resolved in order for a multimodal version of FDG (or any other grammatical model) to be fully operational.

4.1 How systematic is systematic?

FDG explicitly limits its analyses to those semantic and pragmatic elements that are systematically reflected in the morphosyntax or phonology (Hengeveld & Mackenzie 2008: 38-39). When focusing on the verbal channel, this working definition is already open to debate (compare, for instance, interpretations by Hengeveld & Mackenzie 2014 and Connolly 2014). When examining co-verbal behaviors from a grammatical perspective, however, a dichotomous take on systematicity appears even more problematic. As mentioned before, co-verbal behaviors differ vastly in their degree of stability in form and meaning. Whereas emblematic gestures are systematic in much the same way as lexemes, other types of gestural expression are far less crystallized – exhibiting a high level of flexibility in their form and meaning – and only have features in common on a rather abstract level of

---

6 In the analyses above, the terms ‘modifier’ and ‘operator’ are applied to gesture according to their functional relation to the verbal elements in the utterance. For instance, the ‘ten fingers up’ gesture is analyzed as an operator because it is functionally similar to a numeral: it inherently quantifies something. An alternative approach would be to subject gestures to the same grammaticalization criteria as commonly applied to speech in the grammaticalization literature (e.g., desemanticization). Although some gestures may qualify as grammaticalized according to these principles (e.g., the interpersonal instead of deictic use of pointing gestures; see Schoonjans 2014), the vast majority of gestures would under these conditions be considered ‘lexical’.
analysis (Cienki 2005; Wilcox 2004). A major challenge for a multimodal FDG is therefore to adopt a more gradient notion of systematicity, but at the same time to retain a principled boundary of the scope of interest.

4.2 Multifunctionality

Some grammar-oriented work on gesture gives the impression that the meaning or function of a given gesture is always clear-cut and easily interpretable. In reality, however, most naturalistic gestures do not bear a clear correspondence to a distinct (aspect of) meaning or function. Instead, gestures often perform a range of semantic and interpersonal functions at the same time; a single gesture may for instance concurrently refer to an entity, ascribe a property to it, and assign focal status to the referential act.

Whether this presents an insurmountable problem for FDG remains to be seen. Although FDG has been criticized for its limited capacity to deal with multifunctionality (Anstey 2008), it does not strictly assume one-to-one mappings between form and function. Instead, FDG’s stance on this matter can be assumed to align with that of Dik’s Functional Grammar:

[A] functional explanation of grammatical phenomena will typically not be based on an assumption of simple form-function correlations, but will instead involve a network of interacting requirements and constraints, each of which may be understood in functional terms itself, but which interact in complex ways and in a certain sense 'compete' for recognition and expression in the final design of linguistic expressions. (Dik 1986: 17-18)

The surface form of an utterance, hence, needs not be fully predictable on the basis of the functions it encodes. Gestural multifunctionality can be understood along the same lines as a natural outcome of the interplay between various types of communicative and contextual constraints. In order to map out their functional organization in relation to these constraints, a better understanding is needed of the level of description of gesture form that is appropriate for identifying these form-function relations.
4.3 Grammar or output?

A related challenge is to comply with FDG’s commitment to limit its analysis to linguistically coded aspects of meaning. It is not always clear where to draw the line between those aspects of gestural expression that reflect some semantic or pragmatic structure, and those that do not (and therefore, in FDG terms, belong to the extra-grammatical Output component).

In particular, it is unclear how FDG may deal with non-discrete aspects of expression. The manual-visual modality allows for expressing meaning in an analogue fashion, e.g. by modulating the location in space where a gesture is performed or the intensity of the movement. Since such modulations are paramount to the way meaning is encoded during multimodal spoken communication, relegating them to the Output component seems too easy a way out. Therefore, clearer criteria need to be formulated for teasing apart formal features of gestures that are grammatically induced and those that are a matter of phonetics.

What makes this issue even more complicated is that not all co-verbal behaviors are necessarily addressee-oriented. Instead, some manual behaviors primarily serve to facilitate speech production, for instance by aiding word retrieval or information packaging (Kita, 2000). Since these are, strictly speaking, functions of gestures, one may argue that they deserve the attention of functionally oriented linguists (at least more than from other schools). However, pursuing to incorporate self-oriented functions of bodily behaviors in a linguistic theory entails a massively broadened (possibly even boundless) conception of grammar. A more fruitful way of dealing with this issue, instead, is to follow clear diagnostics for recognizing speaker-oriented actions from communicative ones. A number of salient physical characteristics of “utterance-dedicated” actions has already been formulated by Kendon (2004:11), among others.

4.4 Same layers or different layers?

Sections 3.1 and 3.2 outlined a close alignment between FDG’s layers and the ways gestures have been studied in the literature. Although novel and informative, this review falls short of answering whether gestural expression is organized according to the exact same structure as the verbal
channel. Discrepancies, in fact, may exist in two directions. First, not all grammatical distinctions relevant to the verbal component of the grammar receive distinctive encoding in gesture. For instance, the fact that particular kinds of grammatical aspect co-occur with longer and more complex gestures does not necessarily imply that the structure of gestural expression respects the aspectual distinctions as recognized by FDG. Conversely, gestures may encode semantic or pragmatic variables that are not verbally expressed. Whereas German does not have grammatical markers for Episodes, the research by Müller et al. (2013) discussed in Section 3.2 provides evidence that Episode-like semantic units can be marked by German speakers’ gestures. This suggests that the Episode layer has psychological reality even for speakers of languages that do not encode it grammatically, but it does not imply that the Episode layer is necessarily relevant for analyzing the verbal side of the grammar.

Notwithstanding this note of caution, it can be concluded from the close alignment between FDG’s layers and the functional characteristics of gestures that a great degree of overlap exists between the structural-functional organization of the two channels. The most likely explanation for this overlap is that gestural expression to a large extent ties into the functional organization of speech. An adaptation in this direction is not utterly surprising, given that gestures rely on speech to be interpretable much more often than vice versa.

4.5 Temporal coordination as a dimension of encoding

Finally, the accommodation of gesture in a grammatical model may revive the tension between pattern-oriented (static) and process-oriented (dynamic) views on grammar (Butler 2008; Mackenzie 2010). Everett (in prep.: 1), for instance, provocingly asserts that “dynamic human gestures interact in complex ways with static human grammars to produce human language” (emphasis in original). A multimodal perspective on language, however, does not necessarily imply such a dualism. Like speech, gesture is produced dynamically and modulated on-line, but can be analyzed and represented in a static, pattern-based fashion (see Section 3.6).

A more substantial challenge follows from the fact that the real-time coordination of speech and gesture influences the meaning of the multimodal utterance as a whole. Speech and gesture are most likely to be interpreted as co-referential or co-ascriptive when produced in close temporal
proximity (Habets et al. 2011; McNeill 1992). McNeill (1992) observes that gestures are typically produced right before the element of speech to which they are most closely related, within a time window of about two seconds, but this characterization is rather rudimentary. Further research is required to enhance our understanding of the way grammatical relations between speech and gesture are manifested as (multimodal) utterances unfold over time. It is for instance an open question whether the points of connection between gesture and FDG’s layers of formulation, outlined in sections 3.1 and 3.2, correlate with differential patterns of temporal alignment between the two channels.

5. Conclusion

In this paper I have argued that a multimodal perspective on language structure can be mutually beneficial for functionally oriented linguists and gesture researchers. Functional grammatical approaches to the structure of multimodal utterances, on the one hand, can provide gesture researchers with analytical tools for studying the speech-gesture relationship in a principled and rigorous manner. Regarding gesture as integral to language, conversely, presents important insights for and challenges to existing grammatical formalisms. Functional Discourse Grammar combines a number of important strengths when seeking to understand the grammar-gesture nexus in greater detail. Most importantly, FDG acknowledges language’s “omnipresent dual functionality” (Hengeveld & Mackenzie 2008: 30) and provides a fine-grained analytical framework for studying points of connection between speech and other modalities.

A novel finding presented in this paper is that gestures have the potential to modulate (or substitute) all of the main layers of pragmatic and semantic organization that FDG recognizes. Gesture’s role in multimodal utterance formation is often comparable to that of a lexical modifier or operator, but gestures can also embody entire Moves in the absence of speech. Although this analysis does not conclusively show that speech and gesture are sensitive to the exact same semantic and pragmatic factors, it corroborates the fruitfulness of an integrative, multimodal model of language structure.

Motivated by these findings and further connections between grammar and gesture research, I have presented a sketch of a multimodal extension of FDG. Various empirical advances are needed in
order to fill in the details of this model, in particular with respect to gesture’s variably systematic, multifunctional and dynamic nature. In order to make the proposed synergy between grammar research and gesture studies fully appreciable, it is essential for linguists to conceive these issues as relevant for grammatical theory. Expansion of the grammar proper is inevitable, one can conclude, at least for those linguists who wish to stay faithful to the nature of actual, everyday language use.

Acknowledgements
* The research reported on in this article was funded by a ‘PhDs in the Humanities’ grant from the Netherlands Scientific Organization (NWO). The author is grateful to Alan Cienki, Riccardo Giomi, Mike Hannay, Lachlan Mackenzie and two anonymous reviewers for their stimulating feedback on an earlier draft. The author also owes thanks to Stefan Kopp and Kirsten Bergmann for granting access to the SaGA corpus.

References


Okrent, Arika. 2002. A modality-free notion of gesture and how it can help us with the morpheme vs. gesture question in sign language linguistics (or at least give us some criteria to work with). In Richard P. Meier, Kearsy Cormier & David Quinto-Pozos (eds.), Modality and Structure in Signed and Spoken Languages, 175-198. Cambridge: Cambridge University Press.


