

Chapter 1. Introduction

1.1 Background: Locating the research

When we speak, we not only produce a sequence of words, but also sometimes make bodily movements which are called gestures. A question then arises as to whether, and if so, how, gesture relates to language. Since many have considered gestures to be paralinguistic behavior, they remained outside the scope of linguistic study for a long time. However, since the 1980s, evidence has been accumulating suggesting that gesture is closely linked with speech, such as the observations that a) gesture and speech are synchronized temporally, semantically and/or pragmatically (McNeill 1992: 23); b) gesture and speech are affected in similar ways by aphasia (McNeill 1992: 24); c) gesture and speech develop in parallel ways in children (McNeill 1992: 24). In this light, what might seem like an obvious difference between gesture and speech is not as straightforward as previously suspected, which furthermore suggests that gestures should be considered in linguistic research.

This integrated view of gesture and speech is currently held by many scholars, most notably by Adam Kendon and David McNeill – who many consider to be the two founding fathers of the field of gesture studies. Although both Kendon’s and McNeill’s positions maintain an integrated view on speech and gesture (see Ladewig 2012 for an overview of the development of this integrated view), they differ in their focus. Kendon’s paradigm has a larger focus on the interactive functions of gesture in relation to speech (Kendon 1980, 1995, 2004), whereas McNeill’s paradigm is primarily concerned with the cognitive process of gesture production in relation to speech production (McNeill 1985, 1992, 2000b, 2005). This thesis is more concerned with questions related to McNeill’s paradigm.

Although there are various linguistic approaches following McNeill’s paradigm, such as systematic functional linguistics (Martinec 2000, 2004; Muntigl 2004) and formalism-based, generative linguistics, e.g., Constraint Grammar (Alahverdzhieva 2013; Alahverdzhieva & Lascarides 2010) and Lexical Functional Grammar (Giorgolo et al. 2011), the cognitive linguistic approach to gesture stands out among them. Importantly, cognitive linguistics has more possibilities to embrace gesture studies than many other approaches of linguistics do (Kok & Cienki 2016). Consideration of gesture could be a considerable benefit for cognitive linguistics in a number of ways. First, a usage-based approach to language research in cognitive linguistics could not be seen as truly adequate without consideration of gesture, since gesture is ubiquitous in language use. Langacker takes the straightforward position that “any

aspect of a usage event, or even a sequence of usage events in a discourse, is capable of emerging as a linguistic unit, should it be a recurrent commonality” (2001: 146), and that the phonological pole of a symbolic unit could include “not only sounds but also gestures and ...” (2008a: 15). Another argument is that it could benefit cognitive linguistics because cognitive linguistics emphasizes the role of meaning, and gesture is integral to meaning (McNeill 1992, 2005). As Steen and Turner (2013: 267) put it in relation to studying television news as data,

Cognitive linguists routinely study basic mental operations and phenomena that are not exclusive to language but that are deployed in language and leave their mark on its structure: mental space phenomena, conceptual integration, categorization, image-schematic structuring and transformation, fictive motion, force dynamics, viewpoint phenomena, scanning... Since the news deploys other modalities than speech and text, it is an obvious project to look for the ways in which these basic mental operations and phenomena are deployed in those other modalities.

Therefore, a truly adequate cognitive model of semantics or grammar could and should consider both verbal and gestural modalities (see also Cienki 2010, 2013a, 2015, 2016).

The relation of gesture to certain areas in cognitive semantics was established earlier, including conceptual metaphor theory – the earliest point of interaction between gesture studies and linguistic research (Calbris 2008; Cienki 1998; Cienki 2013b; Cienki & Müller 2008; Chui 2011, 2017; Langacker 2008b; Sweetser 1998; Webb 1996) – and then metonymy (Mittelberg 2006; Mittelberg & Waugh 2009), image schemas (Cienki 2005, Cienki 2013c), and mental spaces and blending (Liddell 2009; Parrill & Sweetser 2004; Sweetser 2006). Studies concerning relations between gesture and cognitive approaches to grammar followed, and they are receiving increasing interest, such as Cienki (2015), Kok and Cienki (2016), Schoonjans (2014, 2017), Wilcox (2004), Wilcox and Xavier (2013), and Zima (2014, 2017). However, this field (that is, grammar and gesture) is still rather young. A fundamental question, whose answer still remains open, is whether, and if so, in what ways, there is an interaction between gesture and grammar. This thesis aims to elaborate on this issue and contribute to answering the question.

In order to address the question, it is necessary to empirically study gestures accompanying fundamental grammatical constructions in the language system. However, as Zima and Bergs (2017: 1) point out, “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”. More specifically, previous

studies on the relation between gesture and grammar have been mainly concerned with a limited number of specific motion constructions (Cánovas & Valenzuela 2017; Zima 2014, 2017) and discourse constructions in relation to gestures (Bressem & Müller 2017; Jehoul et al. 2017; Lanwer 2017; Mittelberg 2017; Schoonjans 2014, 2017; Stec et al. 2015), but for the most part, they have not considered the most basic and frequently used grammatical constructions in speech, that is, transitive, intransitive, and copular constructions. If these basic and frequently used constructions are not addressed, a deeper understanding of grammatical constructions in relation to gesture will remain elusive. The present thesis thus aims to bridge this gap by taking these constructions as a starting point.

Moreover, given that, from the perspective of cognitive linguistics, the semantics of constructions refers to the construal of conceptual content – an event (this is discussed in terms of the “semantic frame” in Fillmore 1982 or the “domain” in Langacker 1987, 1990, 2008a) – it is necessary to tease apart the means of construal afforded by¹ constructions from the related properties of events in their relations to the accompanying gestures. However, to my knowledge, few studies have addressed this issue, although there have been studies on either event properties in relation to the accompanying gestures (Parrill 2010) or on grammatical constructions in relation to the accompanying gestures (Duncan 2002; McNeill 2003; Parrill 2010; Zima 2014, 2017). This thesis aims to take a step further by investigating the relation between gesture and the means of construal afforded by constructions versus the related properties of events. To this end, the thesis considers groups of constructional alternations which involve the same events but different constructions for describing them.

1.2 Objectives and scope

The overall aim of this thesis is to investigate the multimodality of constructions. The general research question can be formulated as follows: Is there a relation between gesture and grammatical constructions in spoken language, and can these different ontological categories together constitute multimodal constructions? To address this question, I examine to what extent gestures correlate with the means by which speakers construe events², as evidenced by how they encode them in grammatical

¹ Note that the term “afford” means “allow for” or “facilitate”, but not “determine”. A construction affording a certain means of event construal means that the choice of a construction allows a speaker (and indeed also a hearer) to construe an event in a certain way.

² Note that generally speaking, I use the term “event” in a broad sense to include both dynamic activities and stative situations.

constructions in speech, and to what degree these multimodal constructions are conventionalized, as evidenced by how often the gestures are co-expressive with the constructions in spoken language.

Four more specific objectives can be established. The first one is to provide a discussion on theoretical motivations for considering multimodal constructions within the framework of construction grammars. The second one is to empirically explore how and to what degree the kinds of construal afforded by the most basic and frequently used constructions can be reflected in the accompanying gestures. The third one is to ascertain the relations between gestures and the means of construal afforded by these constructions (profiling different segments of an “action chain”) versus certain properties of frames/events evoked in these constructions. The fourth objective aims to extend the third one: it aims to determine whether a different type of construal afforded by a different group of constructions (profiling different participants of a certain segment of an “action chain”) can be reflected in the accompanying gestures, thus revealing whether and how gestural use is sensitive to various types of construal afforded by various constructions.

It is worth pointing out the scope of this dissertation. Firstly, the dissertation primarily concerns manual gestures,³ in particular: co-speech manual gestures. Other types of body movement such as shoulder shrugs, head nods, or eye gaze will be mentioned when the previous studies on multimodality of constructions are reviewed, but they are not considered in the empirical studies of this dissertation. This choice is motivated by the fact that manual movements can represent objects or events in a way that is much more feasible than non-manual movements (Streeck 2009b: Ch.3); similarly, manual gestures not concurrent with speech — typically emblems, pantomime, and speech-linked gestures (e.g., *she’s like* + gesture) — will be referred to in the theoretical discussions and in the literature reviews, but they fall outside the scope of the empirical chapters. This is motivated by the fact that co-speech manual gestures are the most frequent type of gesture, and that they are believed to arise from dynamic interaction between thinking and speaking in real time, and thus to involve many spontaneous forms/variants and usages (McNeill 2005: 5), which makes them an appropriate object of a study to learn about the relation between gesture and thought. Secondly, the dissertation draws upon cognitive approaches to grammar, including both cognitive grammar (Langacker 1990, 2008a) and construction grammars (Goldberg 1995; Croft 2001, et al.). These are simply referred to as construction grammars in this dissertation, since cognitive

³ In practice, this thesis deals with *bimodal* expressions – gestural-speech co-expressions, but they are simply called *multimodal* expressions throughout the thesis, as is usually done in gestural works.

grammar can be seen as a branch of construction grammar in general (Langacker 2005); this is also driven by the fact that both branches of construction grammar share basic tenets, which will be specified more closely in Chapter 2.

1.3 Overview of the dissertation

This dissertation is composed of both theoretical and empirical chapters, which are structured as follows.

Chapter 2 provides the theoretical motivations for the multimodality of constructions, drawing upon basic tenets in construction grammars. It indicates how it is necessary to consider the possible role of gesture in the concept of “construction” and how characteristics of constructions and features of gestures are compatible with each other in theory. This provides a theoretical motivation for multimodal constructions. **Chapter 3** reviews what has been done on the topic of the multimodality of constructions and posits general research questions to be pursued and constructions to be considered in this thesis – including the basic clausal constructions with different types of transitivity and constructional alternations. **Chapter 4** offers an introduction to these basic clausal constructions and constructional alternations. The means of construal afforded by these constructions and the semantic frames evoked by them are discussed in this chapter. After this, more specific research questions in relation to these semantic aspects are raised, which will be addressed in the coming chapters.

As a first empirical chapter, **Chapter 5** explores gesture use (the frequency of gestures produced, the Modes of Representation [MoR] of representational gestures, and the movement qualities of discourse-related gestures) with respect to basic clausal constructions, including high-transitive, low-transitive, intransitive, and copular constructions. This could provide a general sense of how different kinds of construal of events, as expressed verbally, are reflected in the accompanying gestures.

Chapter 6 investigates gesture use (the frequency of representational gestures produced and the Modes of Representation [MoR] of representational gestures) with respect to transitive and intransitive constructions involving the caused change of states or locations (that is, the causative-inchoative alternation). It asks whether the differences found in gestures together with high-transitive and intransitive clauses in Chapter 5 are correlated with the type of construction (transitivity) or simply with the motor-spatial properties of events. It reveals the ways in which gestures relate to the means of construal afforded by transitive and intransitive constructions and to the motor-spatial properties of events respectively.

Chapter 7 is concerned with how high-transitive/intransitive and

low-transitive/copular constructions evoking tool activities relate to gesture use (the frequency of representational gestures produced, the Modes of Representation [MoR] of representational gestures, and properties of Acting gestures). It asks whether the differences found in gestures together with these constructions in Chapter 5 correlate with the type of construction (transitivity), or simply with the action-affording property of referents involved. This chapter establishes the means by which gestures correlate with the type of construction (type of transitivity) and with the action-affording property of referents involved.

Extending Chapters 6 and 7, **Chapter 8** addresses whether gestures (the frequency of representational gestures produced, the Modes of Representation [MoR] of representational gestures, and properties of Acting gestures) differ with regard to a different type of constructional alternation: two constructions which involve the same events (that is, transfer events) but different syntactic encodings (that is, the double object construction and the related prepositional object construction). This type of constructional alternation is called the dative alternation. Since these constructions involve more similar syntactic forms (that is, the same syntactic transitivity) and more similar means of construal than the above constructions do, an examination of this alternation in relation to gesture would be of help in understanding how far the relation of gesture to means of construal afforded by constructions and semantic frames evoked in these constructions could go.

Chapter 9 summarizes the main findings, discusses their implications for the interaction between gesture and grammar and for the notion of multimodal grammar, and considers some future directions.