Multimodality of Transitive, Intransitive and Copular Constructions in Spoken Language

Suwei Wu

School of Foreign Languages, China University of Petroleum, Beijing, China

Abstract   In line with the multimodal view of language, some studies found a relation between gesture and a fundamental grammatical category - transitivity. However, previous studies have simply employed the data of elicited narratives, which were unnatural and not interactive, and they were restricted to a few particular transitive and intransitive events. Against this background, this study employs more natural and interactive data – conversations – and considers a larger variety of grammatical constructions which are basic and frequently used in spoken language, including high-transitive, low-transitive, intransitive and copular constructions. Specifically, using a large amount of conversational data from the Red Hen database, this study examines the use of gestures accompanying the four constructions, and the question to what extent the gestures preferred relate to the ways of conceptualizing these constructions. Results indicate that the use of gestures accompanying the four constructions shows distinctions in the ways in which speakers conceptualize these constructions. These results suggest that these constructions seem to be multimodal in nature, which provides further empirical support for the multimodal stance that gesture is part of language.

Keywords   Gesture, Modes of Representation, High-transitive Construction, Low-transitive Construction, Intransitive, Copular Construction

1. Introduction

When we speak, we not only produce a sequence of words, but also sometimes make bodily movements which are called gestures. In line with the multimodal view of language that gesture and speech come from a same origin or that gesture is part of language [1], there is an increasing number of studies investigating what grammatical categories or constructions in spoken language are coordinated with gesture use, with a particular focus on representational gestures, e.g. Duncan [2], McNeill [3], Parrill [4] and Cienki and Iriskhanova [5].

Among various topics on the relation between representational gesture and grammar, the relation between gesture and the grammatical category of *transitivity* is noteworthy, since this grammatical category in fundamental in the language system. Some studies found that gesture use relate to the use of transitive and intransitive constructions. Using a small dataset (i.e. four narratives of a single stimulus) in American English, McNeill [1] first examined how transitivity in speech relates to the choice of character viewpoint gestures (CVPT) – in which the speaker imitates some action such as the activity of walking – or observer viewpoint gestures (OVPT) – in which one moves one's hands to depict a scene as viewed, such as tracing the path that someone walked along. It is found that transitive and intransitive utterances tend to be accompanied by CVPT and OVPT gestures respectively. Afterwards, Parrill [4] used a larger dataset of narrative language in American English (69 narrations) to verify McNeill's claim based on. In addition, their findings are also supported by Beattie and Shovelton [6], who found that CVPT gestures were more likely to evoke transitive descriptions, while OVPT gestures were more likely to evoke intransitive ones.

However, previous studies have focused on transitive and intransitive constructions, without considering copular constructions. As part of the grammatical category of transitivity, these constructions are found to be much more frequently used in conversations than transitive and intransitive constructions [7, 24]. If these frequently used constructions are not addressed, a deeper understanding of basic grammatical constructions in grammar in relation to gesture will remain elusive. In addition, existing studies have concentrated on high-transitive constructions (that is, those with more dynamic activities), ignoring low-transitive constructions (that is, those with more static activities, including mental activity, perception, and other static relations), although the two types of transitive

constructions are clearly different and are furthermore likely to be accompanied by different gestures. It is thus necessary to consider low-transitive constructions in relation to gesture as well to obtain a more complete picture of the relation between transitive constructions and gesture. Another point is that existing studies draw their conclusions from the data of elicited narratives in experiments. It thus remains unclear whether these conclusions could also be applied to conversations, given that unlike elicited narratives, conversations involve interactions between speakers and listeners, where speakers have more time pressure to speak and gesture.

Taking the above considerations together, the present study investigates whether and how representational gestures relate to four basic and frequently used constructions in conversational English, including high-transitive constructions, intransitive constructions, low-transitive constructions, and copular constructions. Given that in cognitive linguistics grammatical constructions reflect speakers' construal of events, rather than simply the events in the world, specific research questions to be addressed in this study can be formulated as follows:

1) In which ways do speakers make representational gestures with respect to the four basic constructions in interviews?
2) This will lead to a discussion of the following question: To what extent could the above gestural behaviors relate to the ways in which speakers may be conceptualizing these four constructions?

Addressing these questions can shed further light on the ways in which gesture and verbal constructions interact. Specifically, the present paper considers the so-called gestural Modes of Representation [8, 9], presented in section 3.3, which consist of Acting with object (miming an object-oriented action such as placing activity), Acting only (miming a non-object-oriented action such as walking activity), Tracing (tracing an outline of a form), Molding (molding or holding a shape in the air) and Embodying (stand for some entity by Embodying it).These gestural Modes of Representation are adopted since they could nicely demonstrate the dynamicity and the complexity (i.e. external causation) of the action that a speaker wants to express in gesture.

In the following, I first give a brief introduction to the four types of constructions and then present the method and results for the research question. I end by discussing the implications of the present study.

2. High-transitive, Low-transitive, Intransitive and Copular Constructions

In cognitive linguistics, meaning is equal to conceptualization [10, 11]. In other words, the choice of grammatical constructions depends on speakers' construal of events, rather than simply on the events in the world. It follows that high-transitive, intransitive, low-transitive and copular constructions in speech afford different means of event construal.

First, by using a high-transitive construction (e.g. *Floyd broke the glass*), a speaker is prototypically profiling a whole event – involving an external causation between two participants and a change it brings about, as shown in Figure 1. For example, in the case of *Floyd broke the glass*, *Floyd* is an "agent" and he exerted some force, which caused the glass to be broken. That is, energy is transferred physically from the "agent" *he* to the "patient" *glass*.
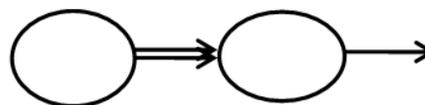

Figure 1.   Schema for high-transitive construction (adapted from Langacker [11], p. 385) (circles represent participants; double arrows represent energy transfer between participants; single arrow represents the change that the energy transfer brings about)

Second, by using an intransitive construction, a speaker is prototypically profiling an entity's change only (either an agentive, e.g., *I walked to the window,* or a non-agentive one, e.g. *the door opened*), without its external causation. This can be indicated in the following figure (see Figure 2).

Figure 2.   Schema for intransitive constructions in general

Third, by using a copular construction (e.g., *we are on the receiving line*), a speaker is prototypically profiling the endpoint of an activity or a static situation, namely an entity with a property or location. To put it differently, it can be said that, by using a copular construction, a speaker simply selects a certain entity in a series of events to profile; that is, only a certain point of an action chain is profiled. The characterization of this construction can be represented in the following schema (see Figure 3).
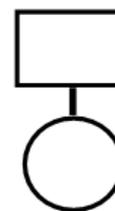

Figure 3.   Schema for copular construction (adapted from Langacker [11], p. 385) (Small square represents attributes or entities; circle represents participants; single line represents a linking relation between the two.)

Last, a low-transitive construction involves a profile of

speakers' mental interaction between two participants in a static activity/situation. It can be represented as follows:
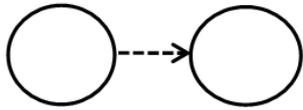

Figure 4. Schema for low-transitive construction (adapted from Langacker [11], p. 385) (the single dashed arrow represents the mental interaction between participants)

All the above has shown that the world consists of a series of interactions between participants. The choice of different constructions is seen as reflecting the fact that we conceptualize the world in different ways. To summarize, speakers' conceptualizations of the four types of transitivity in speech differ in terms of whether they involve a dynamic activity (high-transitive and intransitive clauses) or a static activity or situation (low-transitive and copular clauses); in the case of a dynamic activity, whether it includes an external causation (high-transitive clauses) or not (intransitive clauses); in the case of a static activity, whether it involves a mental contact between two participants (low-transitive clauses) or only a static entity with a certain property or location (copular clauses).

3. Method

3.1. Database

This study is based on a database called the UCLA NewsScape Archive. It is being developed by the Department of Communication at UCLA and the Distributed Little Red Hen Lab, directed by Francis Steen and Mark Turner. The database contains more than 250,000 hours of television and video programs as well as closed captioning, transcripts, and on-screen texts for them from 2005 to the present. The database covers a variety of programs, such as news broadcasts, talk show programs, and presidential debates. It can be accessed with permission through an online search engine at http://newsscape.library.ucla.edu/.

For the current research, talk show programs in this database were used. This was motivated by the fact that they involve conversations, and that the speakers' hands are mostly visible on the screen, which is different from the situation with news broadcasts (that often only show "talking heads"). Particularly, due to the present author's interest, *The Ellen DeGeneres Show* was used. The Ellen DeGeneres Show usually consists of the following portions: opening performances, monologues of the host, interviews (conversations), and film clips for promotion or other advertisements, as well as on-set product promotions at the end of the show. Of these, the interview part is the focus of this study, since the current research specifically focuses on conversations.

The current corpus contains 20 interviews, consisting of the host – Ellen DeGeneres – and 20 different guests, who were actors/actresses, singers, writers, politicians, or other kinds of celebrities. Each interview spanned approximately five minutes, and the total corpus involves around 100 minutes of video.

3.2. Construction Type Coding

Each simple clause in the corpus was coded for its type of transitivity: *high-transitive, low-transitive, intransitive, copular,* or *others*. As mentioned above, **high-transitive** clauses refer to transitive clauses which depict dynamic events, mainly involving events of dynamic actions (that is, they usually involve physical or metaphorical energy transfer between two participants), such as a placing/taking/bringing activity and an act of verbal communication such as saying/talking/telling. **Low-transitive** clauses refer to transitive clauses which depict (relatively) static activities/situations, mainly involving activities of: cognition, such as the activity of thinking; of perception, such as activities of seeing or hearing; of emotion, such as the activity of scaring, as well as involving static relations, such as a relation of possession. It should be noted that clauses with null argument instantiations (no objects) which refer to activities, such as *do you cook, Have you ever filmed in a museum,* were coded as intransitive. The motivation is that these events were formally encoded in intransitive clauses, and semantically they could be understood as activities which did not emphasize the object, as with dancing activities [12]. Clauses with copular *be* (in any of its conjugated forms) and linking verbs were coded as "copular". Those with lexical intransitive verbs were coded as **"intransitive"**. "**Others**" refer to clauses with null argument instantiations which usually refer to cognition, perception, and mental activities. Examples are *I know, I see,* and *I understand*, which involved no objects either right before or after the verb. The reason for this decision is that these clauses were formally intransitive but semantically transitive, since they were difficult to be understood as non-object-oriented activities (that is, different from the above activities in intransitive clauses). Thus they were not seen as transitive or intransitive clauses.

As a result, the corpus consists of 292 high-transitive clauses, 299 low-transitive clauses, 254 intransitive clauses, and 532 copular clauses, as well as 26 others.

3.3. Gesture Coding

3.3.1. Representational Gesture Identification

All simple clauses with the four types of constructions (N=1377) were annotated as to whether they were temporally and semantically accompanied by

representational gestures or not. A representational gesture refers to manual movement which can represent the physical property of an object or event, such as the size, quantity or shape of an object, or the trajectory movement of an event [13, 14].

Fifty gestures were given to a second coder to obtain inter-coder validity of annotations. Inter-coder reliability was 88%. After a discussion on the cases of disagreement, the first coder re-coded the rest of the data. In the corpus, the total number of representational gestures is 332.

3.3.2. Modes of Representation Coding

All representational gestures coded above were furthermore annotated in terms of their Modes of Representation, comprising the categories of "Acting with object", "Acting only", "Tracing", "Molding", and "Embodying" [8, 9]. A gesture was coded as belonging to the Acting-with-object mode when a speaker moved as if miming an object-oriented activity, such as transferring an object. Technically, the mode involves acting as if with the object. The physical object is not actually present, so it is not a matter of actual object manipulation, but of manipulating the imagined object. A gesture was coded as one belonging to the Acting-only mode when a speaker moved as if miming a non-object-oriented activity, such as dancing. A gesture was coded as belonging to the Tracing mode when a speaker moved his/her hands or fingers as if to trace a line, like the outline of an object or the path of a motion. A gesture was coded as belonging to the Molding mode when a speaker moved as if to mold, touch or feel the shape of an object (see also [15]). Finally, a gesture was coded as an instance of the Embodying mode when a speaker used his/her hands to represent an object. For example, a flat open hand was used to represent a piece of paper or a towel.

Sometimes, the speaker combined more than one mode from those above. In these cases, the order of coding priority was "Acting with object" or "Acting only"> "Tracing"> "Molding" or "Embodying", since this research is interested in whether the dynamic activities can be depicted in gestures or not (e.g., the Tracing or Embodying gestures) and in which ways (that is, the Acting-with-object, Acting-only or Tracing mode). For instance, if the Embodying mode was combined with the Tracing mode in one gesture, it was coded as the Tracing mode.

Fifty representational gestures were given to a second coder. The agreement was 62%, but cases of disagreement were restricted to a few particular types. After establishing an understanding of the differences in annotations, the main coder re-coded the rest of the data. In the corpus, there are 71 Acting-with-object gestures (i.e. gestures belonging to the Acting-with-object mode), 34 Acting-only gestures, 95 Tracing gestures, 82 Molding gestures, and 34 Embodying gestures, as well as 16 cases whose modes were uncertain.

4. Results

This section presents how speakers gestured when they used the four constructions in speech. An overview of the results is shown in Table 1.

Chi-square tests were carried out for all the above gestural Modes of Representation accompanying each type of clauses and the other types of clauses in the corpus. The following significant correlations were obtained. First, Acting-with-object gestures were found to significantly correlate with the high-transitive clauses ($p<0.01$, $\chi^2=23.32$, df=1, phi=0.27). Second, intransitive clauses tend to correlate with the use of Tracing ($p<0.01$, $\chi^2=23.79$, df=1, phi=0.27) and Acting-only gestures ($p<0.01$, $\chi^2=30.64$, df=1, phi=0.31). Third, low-transitive clauses are significantly correlated with Molding gestures ($p<0.01$, $\chi^2=8.45$, df=1, phi=0.03; although the effect size for this is very low). Fourth, Molding ($p<0.01$, $\chi^2=16.11$, df=1, phi=0.05) and Embodying gestures ($p<0.01$, $\chi^2=7.46$, df=1, phi=0.15) were found to significantly correlate with copular clauses.

Table 1. Various gestural Modes of Representation (MoR) with respect to various types of transitivity in speech

Type of construction	Acting with object		Acting only		Tracing		Molding		Embodying		TOTAL	
	N	%	N	%	N	%	N	%	N	%	N	%
High TRAN	38	41	4	4	26	28	20	21	6	6	94	100
INTR	13	13	25	26	48	50	7	7	4	4	97	100
Low TRAN	11	19	3	5	9	16	24	42	10	18	57	100
COPULAR	9	13	2	3	12	18	31	45	14	21	68	100

5. Discussion

This section discusses to what extent the gestures that speakers preferred (as presented in the previous section) relate to the way in which a speaker may be conceptualizing each construction.

5.1. Gestures Occurring with High-transitive Constructions

Acting-with-object gestures were found to significantly correlate with high-transitive clauses, those with dynamic transitive events such as placing, giving, sending and carrying activities. For instance, When a speaker described an activity whereby he put a battery into a microphone, he said, *put it in (the microphone)*, and at the same time he made the gesture as follows: he moved two fists together, as if he were putting a battery, which was held in his right hand, into the microphone, held in his left hand.

Based upon the Gestures-as-Simulated-Action framework [16], multiple choices are available for speakers who made gestures accompanying high-transitive constructions in speech, such as miming the relevant activities, tracing a general movement of the activity or a trajectory of the participant's change involved. The preference for Acting-with-object gestures found above is probably motivated by the hypothesized means by which a speaker may conceptualize a high-transitive construction. On the one hand, Langacker [10, 11] proposes that, by using a high-transitive construction, a speaker profiles both external causation involving two participants (animate beings or inanimate entities) (also referred to as energy or force transfer from one participant to another) and a change it brings about in the affected participant. Transmission of force implies 'relative control' of the former participant over the latter one [17]. On the other hand, Acting-with-object gestures refer to object-oriented activities (that is, orienting towards another object), in which one animate entity manipulates another animate or inanimate entity. It follows that there appears to be a mapping between the gestural representation and the way in which a speaker conceptualizes the high-transitive construction.

It is worth noting that some high-transitive clauses were indeed accompanied by Molding gestures (21%, N=20), although gestures of this mode are not significantly correlated with this construction. These gestures usually occurred with clauses of creation and transformation events. For instance, together with the clause *we made smoothies this morning*, the speaker produced a gesture which molded the shape of the *smoothies' cup*. This might be motivated by the fact that newly created objects usually attract speakers' or listeners' attention, and accordingly they usually become focal topics in the discourse. In other words, the choice of gestural mode is also influenced by event structure or pragmatic functions, besides by the type of construction.

5.2. Gestures Occurring with Intransitive Constructions

Results showed that intransitive clauses tend to correlate with the use of Tracing and Acting-only gestures. Gestures of Tracing gestures, which depict a trajectory movement of an entity, and those of Acting-only gestures, which depicted the manner of a motion activity, expressed non-object-oriented manual activities. That is to say, they only expressed entities' movements, rather than any external causation. These gestures appear to relate to the hypothesized way in which speakers conceptualize intransitive clauses: by using an intransitive clause, a speaker is prototypically profiling an entity's change only, without its external causation.

It is noteworthy that as shown in Table 1 above, speakers preferred to use Tracing gestures more often versus Acting-only gestures to accompany intransitive clauses. It would be of interest to explore which aspects in speech relate to the (different) uses of Acting-only and Tracing gestures. This question will now be examined in more detail, in the hope of gaining a more comprehensive understanding of gestural use with respect to intransitive constructions in spoken language.

Previous studies suggest that the language one speaks influences the gestural representation when depicting motion events [18, 19]. More specifically, McNeill [19] proposes the Manner Modulation Hypothesis, which predicts that speakers of satellite-framed languages, such as English, tend to gesture the path information (i.e. Tracing gestures) together with expressions of motion events to downplay the importance of manner information, since the manner information usually has to be expressed via verbs in such languages even when speakers do not intend to highlight the manner information. Following from this hypothesis, it can be said that more Tracing gestures than Acting-only gestures accompanying intransitive clauses in this corpus might be the result of speakers intending to downplay the manner information in speech. The rest part of the section will test this hypothesis. Specifically, it aims to investigate whether the manner component is expressed in actual speech or not, and furthermore, whether the speech encoding the manner information is usually accompanied by gestures depicting the path information (roughly equal to Tracing gestures), as predicted by the Manner Modulation Hypothesis; if not, I will seek to determine which aspects in speech relate to the use of Tracing and Acting-only gestures.

In order to address the questions proposed above, all intransitive clauses accompanied by Acting-only and Tracing gestures were noted for the presence of manner and path. Path is defined as a description of a change of location, which is usually expressed via prepositional phrases (e.g., into the room) or the inherently directed

motion verbs (e.g., go, come), such as *it came out of nowhere about a year ago*. Manner refers to the internal structure of a motion, such as *she was dancing*. If both manner and path were expressed in speech, this was coded as "combination", such as *she ran to school*. If events in the above intransitive clauses did not contain any motion event information, they were coded as "no motion", such as activities of appearance/disappearance (e.g., appearing), posture activities (e.g., sitting), and body-internal activities (e.g. waving).

An exploration of the data indicates that the most frequent type of intransitive clause accompanied by Acting-only and Tracing gestures in the dataset was the one that expressed "path only" (frequent verbs including *go, come, leave,* and others), which accounts for 49% (35/72). The next most frequent ones were those expressing "manner only" (15% – 11/72; such as dancing with someone), the "combination" (18% – 13/72; such as flying into somewhere), and those with no expression of the motion event (18% – 13/72; such as waving). The results indicate that although English, a satellite-framed language, possesses many manner verbs, speakers in the data frequently used verbs which simply depict the "path only" in actual use, e.g., *you go away and come back*. This is inconsistent with the assumption adopted in the Manner Modulation Hypothesis, as mentioned above.

It is also found that intransitive clauses expressing "manner only" and those not expressing motion events were more likely to be accompanied by Acting-only gestures — 64% & 69%, whereas intransitive clauses expressing "path only" and the "combination" (of manner and path) were more likely to be accompanied by Tracing gestures — 83% & 85%. Interestingly, these results indeed do not support the Manner Modulation Hypothesis, which predicts that speakers tend to use Tracing gestures (that is, path gestures) to downplay the importance of manner. Instead, it suggests that the use of Acting-only and Tracing gestures seems to relate to the encodings of manner and path components in intransitive clauses in English. Taking the number of each type of intransitive clause above into consideration, it can be said that the reason for the greater number of Tracing gestures versus Acting-only gestures occurring with intransitive clauses may simply be that intransitive clauses encoding the path component (that is, clauses with "path only" and "combination" of path and manner; 67% – 48/72) were more frequently used than those encoding the manner information (that is, clauses with "manner only"; 15% – 11/72) in the actual spoken usage here. This explanation is more satisfying than one resorting to a hypothesized need for speakers to modulate the manner in English in general.

5.3. Gestures Occurring with Low-transitive Constructions

Low-transitive clauses, mainly involving activities of

volition (e.g., wanting, thinking), mental states (e.g., liking, forgetting), perceptions (e.g., seeing, hearing), possession (e.g., having) in the corpus, are significantly correlated with Molding gestures. Molding gestures, which usually simply depicted entities, seem to be associated with one means of event construal afforded by low-transitive constructions but not another. On the one hand, the gestures reflect the static properties of these events expressed in low-transitive clauses, that is, no dynamic energy transfer between two participants. On the other hand, these gestures seem not to reflect the following subjective way speakers may have of conceptualizing this type of clause: a subjectively dynamic interaction or mental contact between two participants, as hypothesized in low-transitive clauses. To put it another way, the way of construing low-transitive clauses (that is, the mentally dynamic interaction between two participants) may be less likely to be reflected in the mode of gestural representation, compared with the other kinds of construal of events discussed above (i.e. the dynamicity).

It is worth noting that some low-transitive clauses were indeed accompanied by Acting-with-object gestures (19.3%, N=11) and Tracing gestures (15.79%, N=9), although gestures of these modes are not significantly correlated with this construction. For instance, when a speaker described a situation involving celebrities' stylists fighting for dresses, so that the celebrities can wear them in a ceremony, she said *I want this and this*, and at the same time she made a gesture as shown in Figure 5. In speech, there existed no actual energy interaction between two participants – *I* and clothes – but a mental interaction between the two participants is claimed to exist [10, 11]. In the gesture produced, the speaker moved and held her fists first on her left and then on her right, as if she were grabbing some objects (that is, referred to as a gesture of the Acting-with-object mode). This suggests that the mental desire seems to be reflected in the visibly dynamic grabbing movements.

Figure 5.   Gesture with *I want this*

These gestures (Acting-with/without-object or Tracing gestures) were usually carried out to depict how a person desired something in "desire events" (e.g. *I want this*), how a person thought of something in psychological events (e.g.

she thought that…), or how a person perceived something in perception events (e.g. *I felt stuff*). The use of these gestures seems to suggest that the accompanying gestures could sometimes reflect abstract, mentally dynamic contact in speakers' conceptualizing low-transitive clauses, one means by which speakers may construe low-transitive clauses, as suggested by Langacker [10, 11].

5.4. Gestures Occurring with Copular Constructions

Molding and Embodying gestures were found to significantly correlate with copular clauses. As noted in Section 2, the choice of dynamic activities or static situations expressed in speech is not simply a direct reflection of reality in the physical world. Instead, it manifests speakers' construal of reality. More specifically, by using a copular clause, a speaker prototypically profiles an endpoint of an activity or a static situation – an entity with a property or location – rather than any dynamic process related to it. Molding and Embodying gestures, rather than Acting (with/without object) or Tracing gestures, were found to preferably accompany copular constructions, which suggests that the modes of gestural representation preferred tend to reflect the way in which speakers conceptualize this construction (that is, a profile of a static situation of an entity, rather than any dynamic process related to the entity). Given that existing studies did not consider copular constructions – constructions that were frequently used in the spoken language -- in relation to the accompanying gestures, the present results provide a more comprehensive understanding of the relation between gesture use and the type of constructions in speech.

To summarize, all the above analyses show that the modes of gestural representation preferably relate to the means of event construal afforded by the four constructions in speech, rather than simply the events in reality. However, gestures accompanying low-transitive clauses display a slightly different pattern: the coordination between stasis of events and the gestural modes bypasses a correlation between the subjective conceptualization involved (i.e. dynamic mental contact) and gesture.

6. Conclusions

The current study has investigated a) the use of gestures accompanying the four basic and frequently used constructions in spoken language – high-transitive, low-transitive, intransitive and copular constructions, and b) the question to what extent the gestures used correlated with the ways of conceptualizing these constructions in conversations (that the gesture accompanied) as theorized in the literature. Through an examination of a corpus of talk show interviews in English, it was found that there is a correlation between gesture use and the use of grammatical constructions in speech and that the use of gestures

accompanying the four types of constructions shows distinctions in the ways in which speakers conceptualize these four types of constructions. The results concerning high-transitive and intransitive constructions in relation to gesture use are consistent with McNeill's [1] and Parrill's [4] findings. The results as for low-transitive and copular constructions in relation to gesture use are further informative about the ways in which gesture could express different kinds of event construal afforded by the basic and frequently used constructions in conversations. These results together suggest that the four types of constructions are closely linked with gesture use or that these constructions are multimodal in nature. Establishing the multimodality of these constructions provides more empirical support for the multimodal stance that gesture and speech come from the same "idea unit" or that gestures should be seen as part of language [1, 18]. This furthermore challenges "traditional" cognitive linguistics or linguistics in general, which mostly focus on written text or verbal speech. Instead, it calls for a multimodal approach to language and communication in cognitive linguistics or linguistics in general (see also [20-22].

The above correlations between gesture and conceptualization can also offer a window onto the understanding of these means by which speakers conceptualize the four constructions, as hypothesized by Langacker [10, 11]. Whereas the proposed difference in conceptualizing transitive and intransitive constructions (that is, with or without a profile of an external causation) has gained support in previous imaging studies [23], the above finding in terms of the gestural difference apparently provides a different tool to verify this proposal. In addition, since the proposed mental contact between two participants in speakers' construal of low-transitive clauses (that is, construal of physically 'static' events) seems more elusive and more difficult to perceive than the other two means of event construal, the use of gesture which is related to this means of construal (Acting-with-object gestures) has a particularly clear, more important implication. The gestural data show that the mental contact in the construal, proposed for this construction, appears to have cognitive reality.

Several extensions of the research are possible. One fruitful avenue for further research would be to investigate how sub-types of each kind of transitivity relate to gesture. For instance, provided that each type of transitivity is a multi-dimensional phenomenon, the high-transitive construction in this research can be further classified in terms of the parameters of agency (e.g., *she broke the window* vs. *the wind broke the window*), intentionality of the subject (e.g., *she baked a cake* vs. *she dropped the cake on the floor*), affectedness of the object (e.g., *she put it on the table* vs. *she's holding it in her hand*), and others (a list of parameters which have been used to categorize transitivity can be found in [24]). It would be worthwhile to examine to what extent gesture correlates with these sub-types of transitivity, so that more insight can be gained

concerning the degree of, and the sites of, connection between spoken language and gesture – the issue at the heart of research on the multimodality of grammar. In addition, whereas this thesis has dealt with the semantics (i.e. means of construal) of grammatical constructions in relation to gestural use, it would be of interest for future studies to investigate how gestures relate to pragmatic functions of these constructions. Since each construction is situated in a discourse context, the semantics of constructions could be intertwined with their pragmatics in discourse. Hence, including the relation of gestures to the pragmatics of constructions would help to establish a more comprehensive picture of the multimodality of constructions.

Acknowledgements

REFERENCES

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

[2] Duncan, S. D. Gesture, verb aspect, and the nature of iconic imagery in natural discourse. *Gesture*, Vol. 2, No. 2, 183–206, 2002.

[3] McNeill, D. Aspects of aspect. *Gesture,* Vol. *3,* No. 1, 1–17, 2003.

[4] Parrill, F. Viewpoint in speech–gesture integration: Linguistic structure, discourse structure, and event structure. *Language and Cognitive Processes*, Vol. 25, No. 5, 650–668, 2010.

[5] Cienki, A., & Iriskhanova, O. *Aspectuality across languages: Event construal in speech and gesture*. Amsterdam: John Benjamins, 2018.

[6] Beattie, Geoffrey, Heather Shovelton. What properties of talk are associated with the generation of spontaneous iconic hand gestures?.*British Journal of Social Psychology,* Vol.41, 403-417, 2002.

[7] Thompson, S. A., & Hopper, P. J. Transitivity, clause structure, and argument structure: Evidence from conversation. In J. L. Bybee & P. J. Hopper (Eds.), *Frequency and the emergence of linguistic structure*. Amsterdam: John Benjamins, 27–60, 2001.

[8] Müller, C. Iconicity and gesture. In S. Santi, I. Guaïtella, C. Cavé & G. Konopczynski (Eds.), *Oralité et gestualité: Communication Multimodale, Interaction* (pp.). Paris: L' Harmattan, 321–328, 1998.

[9] Müller, C. (2014). Gestural modes of representation as techniques of depiction. In C. Müller, A. Cienki, E. Fricke, S. Ladewig, D. McNeill & J. Bressem (Eds.), *Body –

language – communication: An international handbook on multimodality in human interaction*, *Vol. 2*. Berlin and Boston: De Gruyter Mouton, 1687–1702, 2014.

[10] Langacker, R. W. *Concept, image, and symbol: The cognitive basis of grammar*. Berlin: Mouton de Gruyter, 1990.

[11] Langacker, R. W. *Cognitive Grammar: A basic introduction*. Oxford: Oxford University Press, 2008.

[12] Newman, J., & Rice, S. Transitivity schemas of English EAT and DRINK in the BNC. In S. Th. Gries & A. Stefanowitsch (Eds.), *Corpora in Cognitive Linguistics: Corpus-based approaches to syntax and lexis*. Berlin: Mouton de Gruyter, 225–260, 2006.

[13] Kendon, A. *Gesture: Visible action as utterance*. Cambridge: Cambridge University Press, 2004.

[14] Kita, S. How representational gestures help speaking. In D. McNeill (Ed.), *Language and gesture*. Cambridge: Cambridge University Press, 162–185, 2000.

[15] Streeck, J. Depicting by gesture. *Gesture*, Vol. *8,* No. 3, 285–301, 2008.

[16] Hostetter, A. B., & Alibali, M. W. Visible embodiment: Gestures as simulated action. *Psychonomic Bulletin & Review*, Vol. 15, No. 3, 495–514, 2008.

[17] Croft, W. A. *Verbs: Aspect and causal structure*. Oxford: Oxford University Press, 2012.

[18] Kita, S., & Özyürek, A. What does cross-linguistic variation in semantic coordination of speech and gesture reveal? Evidence for an interface representation of spatial thinking and speaking. *Journal of Memory and Language, 48*(1), 16–32, 2003.

[19] McNeill, D. Imagery in motion event descriptions: Gestures as part of thinking-for-speaking in three languages. In A. C. Bailey, K. E. Moore & J. L. Moxley (Eds.), *Proceedings of the Twenty-Third Annual Meeting of the Berkeley Linguistics Society*. Berkeley, CA: Berkeley Linguistics Society, 255–267, 1997.

[20] Cienki, A. Spoken language usage events. *Language and Cognition*, Vol. 7*,* No. 4, 499–514, 2015.

[21] Cienki, A. Cognitive linguistics, gesture studies, and multimodal communication. *Cognitive Linguistics*, Vol. 27, No. 4, 603–618, 2016.

[22] Steen, F., & Turner, M. Multimodal Construction Grammar. In M. Borkent, B. Dancygier & J. A. J. Hinnell (Eds.), *Language and the creative mind*. Stanford, CA: CSLI Publications, 255–274, 2013.

[23] Thompson, C. K., Lange, K. L., Schneider, S. L., & Shapiro, L. P. Agrammatic and non-brain-damaged subjects' verb and verb argument structure production. *Aphasiology*, Vol. 11*,* No. 4-5, 473–490, 1997.

[24] Hopper, P. J., & Thompson, S. A. Transitivity in grammar and discourse. *Language*, Vol. 56, No. 2, 251–299, 1980.