Chapter 5. Basic clausal constructions (transitivity) and gesture

5.1 Introduction
This chapter investigates how speakers make gestures accompanying the most basic and frequently used constructions in a set of conversations in English (that is, high-transitive, intransitive, low-transitive, and copular constructions), and furthermore whether, and if so, how, these gestures relate to the means of event construal afforded by these grammatical constructions. This can shed further light on the ways in which gesture and verbal constructions interact.

As noted in Chapter 4, these constructions with various types of transitivity in speech afford different means of event construal. Specifically, by using a high-transitive construction, a speaker is prototypically profiling a whole event – involving an external causation between two participants and a change it brings about; by using an intransitive construction, a speaker is prototypically profiling an entity’s change only (either an agentive or non-agentive one), without its external causation; by using a copular construction, a speaker is prototypically profiling the endpoint of an activity or a static situation, namely an entity with a property or location; and a low-transitive construction involves a profile of speakers’ mental interaction between two participants in a static activity/situation. To summarize, speakers’ construals of the four types of transitivity in speech differ in terms of whether they involve a dynamic activity (in high-transitive and intransitive clauses) or a static activity/situation (in low-transitive and copular clauses); in the case of a dynamic activity, whether it includes an external causation (in high-transitive clauses) or not (in intransitive clauses); in the case of a static activity/situation, whether it involves a mental contact (i.e. subjective interaction) between two participants (in low-transitive clauses) or only profiles a static entity with some property or location (in copular clauses) (also meaning more or less complex in terms of the length of the action chain involved in various constructions).

Furthermore, this chapter considers various types of gesture, including representational, deictic, and discourse-related gestures, driven by the following motivations. As mentioned in Chapter 3, previous research on gesture and grammar has predominantly concentrated on representational gestures and has largely ignored discourse-related gestures, although discourse-related gestures are usually more frequently used in the conversational data than representational gestures are, as will be shown below in Section 5.3. It thus remains a mystery as to whether discourse-related gestures correlate with grammar in speech or not, or whether
representational and discourse-related gestures derive from the same cognitive origin or not. The present study aims to provide insights into this issue by including representational gestures as well as discourse-related gestures in the research. In addition, deictic gestures will be touched upon, but not examined in detail, given their less frequent use in the corpus, as will be shown below in Section 5.3.

Taking the above considerations together, the present chapter investigates whether and how gestures of various types (including frequencies and forms of representational, deictic, and discourse-related gestures) correlate with four basic and frequently used constructions in conversational English, including high-transitive constructions, intransitive constructions, low-transitive constructions, and copular constructions. In other words, it investigates to what extent the gestural use correlates with the means of event construal afforded by the four constructions, including +/-dynamic, +/-external causation, and +/-mental contact. Specific research questions to be addressed in this chapter can be formulated as follows:

a) How do speakers use the four basic constructions in spoken interviews?
b) How often do speakers make gestures in interviews (in terms of representational gestures, deictic gestures, and discourse-related gestures) with respect to the four basic constructions?
c) In which ways do speakers make representational gestures with respect to the four basic constructions in interviews?
d) In which ways do speakers make discourse-related gestures with respect to the four basic constructions in interviews?
e) 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 (that is, in terms of +/-dynamicty, +/-external causation, and +/-mental contact/subjective interaction)?

5.2 Method

5.2.1 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

Deictic gestures are not examined in this way, given their less frequent use in the corpus, as mentioned above.
presidential debates. It can be accessed with permission through an online search engine at http://newsscape.library.ucla.edu/.\(^{27}\)

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. It is worth noting that the talk show interview part usually contains the following segments: greetings, personal story retelling, and introduction to guests’ films, songs, or books, etc. Also note that talk show interviews usually exhibit discourse features characteristic of institutional discourses, such as purposefulness, institutionalness, host-controlledness, and multiple audience-orientedness; but to some extent they also have discourse features pertaining to casual conversations, such as spontaneousness, non-institutionalness, non-controlledness, interlocutor-orientedness, and message-orientedness (Ilie 2001).

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. It is important to note that the other analyses of gesture use in Sections 5.3 and 5.5 related to 10 videos, while the analyses of gestural modes of representation in Section 5.4 drew upon the whole dataset of 20 videos, since the number of representational gestures in the 10 videos was not sufficient for a quantitative analysis.

5.2.2 Speech coding

5.2.2.1 Clause transcription

Based on the speech transcripts provided by the database,\(^{28}\) utterances in the corpus were segmented into clauses. As noted in Chapter 4, (Langacker 2001), a clause is usually headed by a lexical verb or linking verb together with one or more

\(^{27}\) See https://sites.google.com/site/distributedlittleredhen/home for details.

\(^{28}\) I went through the utterances in the data and modified transcripts which involved errors. Note that the transcripts from the database are mostly correct.
nouns (that is, subject\textsuperscript{29}, or object) as well as some adverbial components (e.g.,
temporal or spatial adverbs). Semantically, it depicts an event of one or more
participants in a setting. Note that both finite and non-finite clauses were included in
the current research.

Utterances with the following types of verbs were not counted as clauses in the
current study:

a) (Semi-)Auxiliary verbs, e.g. be going to, can, have to, should, must, used to,
seem to, appear to, be likely to, try, and manage to, which usually add grammatical
meanings to clauses such as speakers’ attitudes towards (the occurrence likelihood of)
events;

b) Aspect cues like begin, start, keep, end up, and wound up, which only
express aspectual properties of events;

c) Invitation markers like let’s, which are more like modal verbs for invitation;

d) Some discourse markers such as you know, I mean, thank you, and thank
God, which simply serve to organize the discourse and can flexibly appear in various
positions in a clause (Schiffrin 1988: 32).

Furthermore, clauses with analytic causative verbs fall outside the scope of the
present research, e.g. have/make/get/cause somebody to do something. These
clauses usually involve sequences of causally related events without temporal or
spatial contiguity, whereas simple clauses prototypically concern single (complex)
events with temporal and spatial contiguity (Langacker 1990: 212). Given the
complexity of clauses with analytic causatives, clauses of this type (N=15) were
excluded from the present corpus.

Consequently, the corpus is comprised of 1,403 simple clauses for research.

5.2.2.2 Construction type (transitivity type) coding

Each simple clause in the corpus was coded for its type of transitivity: high-transitive,
low-transitive, intransitive, copular, or others (specific introduction to these
categories can be found in Chapter 4). The coding procedure was as follows.

Firstly, if a clause involved a direct object, it was coded as “transitive”; if not, it
was coded as “non-transitive”. It is important to note that phrasal verbs were taken
as wholes. For instance, clauses with transitive phrasal verbs such as X look at the
detail were coded as transitive. Provided that the MacMillan Dictionary online\textsuperscript{30}

\textsuperscript{29} Note that subjects in clauses were omitted sometimes, such as (he went away), came
back and then....

\textsuperscript{30} Website of the dictionary: http://www.macmillandictionary.com/. It is a corpus-based
dictionary, which thus could reflect general usage of these verbs rather than just that of
experts.
could provide information as to whether verbs together with prepositions or particles are phrasal verbs or not, the decision of phrasal verbs was made on the basis of that dictionary. Examples of transitive phrasal verbs included look at, look through, play with, flip through, work on, and others.

Secondly, a transitive clause was furthermore coded in terms of whether it was a “high-transitive” or “low-transitive” clause. High-transitive clauses refer to those 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 those 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.

Next, non-transitive clauses were further coded in terms of whether they were intransitive or copular clauses. Clauses with copular be (in any of its conjugated forms) and linking verbs were coded as “copular”. The other non-transitive cases (those with lexical intransitive verbs) were coded as “intransitive”.

The following types of cases should be noted.

A) 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 (Newman & Rice 2006).

B) Clauses with null argument instantiations which usually refer to cognition, perception, and mental activities were coded as “others”. 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.

5.2.3 Gesture coding

5.2.3.1 Gesture identification
All simple clauses with the four types of transitivity (N=1377) were annotated as to whether they were temporally and/or semantically/functionally accompanied by
gestures or not. Since this study focuses on co-speech manual gestures, emblems (e.g. the Okay sign), and speech-linked gestures (e.g. I was like + gesture) (total N=23) were first excluded from the dataset. This yielded 645 clauses accompanied by co-speech gestures, 614 clauses without gestures, 86 clauses in which hands were invisible on the screen, and 9 clauses accompanied by real actions.

5.2.3.2 Gesture type coding
All co-speech gestures (N=645) were then coded according to gesture types adapted from McNeill (1992) and Kendon (2004), which include the categories of “representational”, “deictic”, and “discourse-related” gestures.

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 (Kendon 2004; Kita 2000; McNeill 1992). A deictic gesture here refers to manual movement which is used to identify an entity or indicate the location of an entity. It can be carried out with one single finger or a whole hand. A discourse-related gesture refers to one which does not depict any substantial semantic content in speech, but instead is used by speakers to display the rhythm of the speech flow, organize the discourse, or initiate a speech act, etc.

Since representational and deictic gestures are more likely to relate to grammar than discourse-related gestures are, given their semantic nature (as seen in their definitions), this thesis pays more attention to representational (and deictic) gestures than to discourse-related gestures. Given that representational gestures seem more frequently used than deictic gestures do (as will be seen in Section 5.3), this thesis pays more attention to representational gestures than to deictic gestures, as mentioned above. This motivates the following two coding hierarchies. First, when a clause involved more than one co-speech gesture, the priority order for coding was: representational > deictic > discourse-related gesture. For instance, if a clause was accompanied by both representational and deictic gestures, the representational one was coded. Second, when a gesture involved multiple functions, the first priority was given to the label “representational gesture”, the second priority to “deictic gesture” and the last one to “discourse-related gesture”. For instance, representational gestures with deictic functions were only coded as representational. Note that Palm-Up-Open-Hand gestures with weak deictic functions were coded as discourse-related gestures rather than deictic ones, since this gestural family served a discursive or communicative function in all cases regardless of the speech context, and it appears difficult to define a clear cut-off point for these gestures as deictic or as discourse-related.

Fifty gestures were given to a second coder to obtain inter-coder validity of
annotations. Inter-coder reliability was 88%. One major type of disagreement concerned the cases where speakers depicted an abstract object by holding two hands facing each other. For instance, when a speaker said “something” in you got to commit to something, she held two hands facing each other, as if holding an object. The speaker usually made a number of similar gestures in the discourse. This gesture can be seen as depicting an activity of the speaker holding the abstract referent “something” one needs to commit to. In this case, it is considered as an instance of representational gestures. It can also be seen as an instance of “catchments”, where a speaker makes two or more similar gestures (or gestures with one or more similar features) to suggest a common discourse theme (McNeill 2005: 116), and thereby it can be seen as an instance of discourse-related gestures. This study simply takes gestures of this type as representational gestures.

After a discussion on the cases of disagreement, the first coder re-coded the rest of the data. The corpus (10 videos) consists of 169 representational gestures, 82 deictic gestures, and 376 discourse-related gesture, as well as 18 uncertain cases. Since the analysis of gestural Modes of Representation concerned 20 videos, all representational gestures (N=163) in the other 10 videos were also identified. Thus, the total number of representational gestures is 332 (in 20 videos).

5.2.3.3 Modes of Representation of representational gesture 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” (examples and images of these categories can be found in Table 5.1) (Müller 1998, 2014). 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 Streeck 2008). 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%. Cases of disagreement mainly concerned the following types.

A) Surprisingly, one fourth of the cases of disagreement occurred with cases referring to bodily activities (such as squatting) produced with the hand shape of fists (the hand shape someone might have when doing these activities). These gestures can be seen as depicting either non-object-oriented bodily activities (i.e. squatting activities) or object-oriented bodily activities (i.e. activities of carrying up objects together with squatting). Given that manual movements made with the fist hand shape seem in these cases to be a by-product of whole body movements, more than being oriented at manipulating objects, these gestures were agreed as depicting non-object-oriented activities, and accordingly they were coded as the Acting-only mode.

B) Another major disagreement concerned the ambiguity between Acting (including the Acting-with-object and Acting-only modes) and Tracing gestures. An example of the first type of ambiguity was where a speaker moved the two flat open hands close to each other when depicting the activity of placing two dogs near each other. Cases of this type were coded as the Tracing mode rather than the Acting-with-object mode, since the speaker moved his flat open hands together as if tracing the general movement of the placement activity or tracing the movement of two dogs, rather than miming the specific placement activity, whereas more specific hand shapes would be required to mime the object-oriented, placement activity. The other type of ambiguity concerned cases where the manner of an intransitive activity was depicted but not mimed. For example, when depicting an activity of someone turning (and then falling), a speaker moved his/her two flat open hands in an arc line (towards the right of the speaker), as shown in Figure 5.1.
In this example, although the speaker depicted the manner of motion (that is, turning) via the manual movement, he was not miming the activity (that is, twisted the whole body as if turning); instead, his hands can be seen as representing the person and the movement of the person was traced by the speaker’s hands. Since the key difference between the Acting and Tracing modes is whether speakers mimed an activity or not, cases of this type were coded as the Tracing mode in this study.

C) The last major type of disagreement involved gestures referring to the height of entities (as will be illustrated in the second type of Embodying in Table 5.1). Usually, the speakers moved one hand to an upper location in space, which sometimes involved a few beats. Since the speakers did not move in a vertical line as if to trace the line towards the highest point, gestures of this type were considered as the Embodying mode, rather than the Tracing mode.

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.

In order to offer a better sense of the general forms of these gestural Modes of Representation (MoR), a summary of the main sub-types of each Mode of Representation (MoR) together with their examples and images is given below.
# Table 5.1 Summary of main sub-types of each gestural Mode of Representation (MoR) and their examples & figures

<table>
<thead>
<tr>
<th>MoR</th>
<th>Main sub-types</th>
<th>Examples</th>
<th>Figures of gestures</th>
</tr>
</thead>
</table>
| Acting with object | **Object-oriented activities**, such as miming activities of handing over something, putting something in a place, or manipulating something. | **Speech:** *He’s trying to unscrew the microphone.*  
**Gesture:** the speaker twisted his hands (one with bent fingers and one with bunched fingers) in opposite directions, as if unscrewing an object. | ![Figures of gestures](image1) |
| Acting only | **Activities of miming running, dancing, or other non-object oriented activities** (more frequent) | **Speech:** *He was running.*  
**Gesture:** the speaker mimed the activity of running. | ![Figures of gestures](image2) |
| Acting only with embodying | | **Speech:** *The cat stretched like a bat.*  
**Gesture:** the speaker stretched her two arms, miming the activity of a bat stretching out | ![Figures of gestures](image3) |
<table>
<thead>
<tr>
<th>Basic clausal constructions (transitivity) and gesture</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Tracing</strong></td>
</tr>
<tr>
<td><strong>Gesture:</strong> the speaker moved his forefinger towards him in a straight line, to depict the path of an object coming towards him.</td>
</tr>
<tr>
<td><strong>Tracing with flat open hand (the most frequent in the corpus)</strong></td>
</tr>
<tr>
<td><strong>Gesture:</strong> the speaker moved her flat open hand (a bit stretched in this case) away from her body, depicting the trajectory of an entity leaving.</td>
</tr>
</tbody>
</table>

Note that this gesture also involves the Embodying mode (in addition to the Tracing mode coded above) due to the fact that the forefinger can be seen as embodying the microphone. However, according to the coding hierarchy discussed above, that is, “Acting with object” or “Acting only” > “Tracing” > “Molding” or “Embodying”, this gesture was only coded as the Tracing mode.
### Cyclic movements

**Speech:** *She thought.*

**Gesture:** the speaker made a cyclic gesture, to show the process of thinking.

### Molding

**Dynamic molding,** in which speakers move their hands as if to shape an object as a sculpture

**Speech:** *The carpet has a Mickey Mouse in it.*

**Gesture:** the speaker moved her two hands (facing each other), as if following the contour of the Mickey Mouse in the carpet.

**Static molding/holding** (Two hands or fingers facing each other) (the most frequent)

**Speech:** *He’s a domestic cat.*

**Gesture:** the speakers held two hands facing each other, as if touching/holding the surface of the cat.

### Embodying

**Speech:** *There is a current.*

**Gesture:** the speaker stretched out her whole arm, using the arm to represent the (long) river.

---

32 This mode can also be referred to as the Holding (Cienki 2013b, 2017a) or Touching mode (Cienki 2017a).
Embodying something with a certain height (frequent)

**Speech:** *Intelligence usually isn’t the first thing (they say).*

**Gesture:** the speaker moved her hand (four fingers flapping down) to the upper right (together with a few small beats in this case), referring to an entity in a high position.

Finger counting (frequent)

**Speech:** *I only have two (cats).*

**Gesture:** the speaker held out two fingers, representing two entities.

5.2.3.4 Form parameters of discourse-related gesture coding

All the discourse-related gestures (N=376) were coded for the tension and size of movements made during their strokes. The motivations for considering these dimensions in gesture are as follows. Firstly, the movement parameter is argued to be most relevant to the dynamics of an event (Wilcox 2004). Secondly, the tension and size of a manual movement appear analogous to the dynamicity and/or complexity (in terms of the length of an action chain, as mentioned before) in the speaker’s construal of various types of transitivity. Moreover, the size of a manual movement has been found to be used by signers to differentiate dynamic processes from static objects in their expressions (Kimmelman 2009). Means of conceptualization of dynamic processes and static situations seem comparable to those of speakers conceptualizing various types of transitivity, that is, in terms of dynamicity and/or complexity (the length of an action chain).

The label *tension* refers to the effort in making a discourse-related gesture. When a manual movement was effortful – usually involving the use of accelerating movement – it was coded as a gesture with tension. Otherwise, it was coded as one
without tension. Note that compared with the preparation and retraction phases in a gesture, all stroke phases were somewhat effortful. Here a stroke phase with tension was coded on the basis of tension involved in various stroke phases before and after the stroke, rather than on the basis of tension in the preparation or retraction phases around the stroke phase.

The label size of a gesture refers to the most proximal joint involved in a stroke articulation. It can be “finger”, “wrist”, “forearm”, or “arm”.

Fifty gestures were given to a second coder for a cross-check. The agreement for gestural tension was 70% and for gestural size it was 80%. Major disagreement on gestural size concerned cases in which a simple wrist rotation (in Palm-Up-Open-Hand gestures) sometimes occurred together with a slight movement of the forearm. The outcome of the discussion was that these cases should simply be seen as gestures produced with the “wrist” rather than the “forearm”.

After a discussion on the cases of disagreement, the first coder re-coded all the remaining data. The corpus contains 134 gestures made with tension and 242 cases without tension. In addition, it contains 7 cases produced with fingers, 100 with wrists, 236 with forearms, and 33 with arms.

5.3 Gestural rates with respect to high-transitive, intransitive, low-transitive, and copular constructions

I begin with an overview of the use of the four constructions in speech and the use of gestures in interaction. After that, I will consider gestural use with respect to each construction in speech, and discuss to what extent the frequencies of gestures used relate to the means of event construal afforded by each construction.

5.3.1 High-transitive, intransitive, low-transitive, and copular constructions in speech

This section presents what the speakers were uttering in speech with respect to the four constructions. The frequencies of the four basic constructions in speech are shown in Table 5.2.
Table 5.2 Distribution of various types of transitivity in speech (count-N & proportion-%)

<table>
<thead>
<tr>
<th>Type of construction</th>
<th>Frequency in the corpus</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
</tr>
<tr>
<td>High TRAN</td>
<td>292</td>
</tr>
<tr>
<td>INTR</td>
<td>254</td>
</tr>
<tr>
<td>Low TRAN</td>
<td>299</td>
</tr>
<tr>
<td>COPULAR</td>
<td>532</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>1377</strong></td>
</tr>
</tbody>
</table>

In general, speakers in the dataset were somewhat more likely in this study to utter clauses with lower transitivity (around 60%), including copular (39%) and low-transitive clauses (22%) than clauses with higher transitivity (around 40%), including high-transitive (21%) and intransitive clauses (18%). Drawing upon the data from talk show interviews, this result is consistent with that of spontaneous conversations in a study by Thompson and Hopper (2001): English speakers were more likely to utter clauses with lower transitivity than those with higher transitivity. The frequent use of copular and low-transitive clauses (clauses with lower transitivity) points to the need for a broader, more comprehensive investigation of gestural use with respect to a greater variety of constructions in communication, beyond studies of gestural use of motion events, which is a frequent subject of research in gesture studies.

5.3.2 An overview of representational, deictic, and discourse-related gestures in interviews

This section presents how the speakers were gesturing in general in the interviews. Table 5.3 displays how frequently the speakers produced gestures in general and in terms of individual gestural categories.
Table 5.3 Frequencies of representational, deictic, and discourse-related gestures in interviews

<table>
<thead>
<tr>
<th>Type of clauses accompanied by gestures</th>
<th>Frequency in the corpus</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
</tr>
<tr>
<td>Clauses with representational gestures</td>
<td>169</td>
</tr>
<tr>
<td>Clauses with deictic gestures</td>
<td>82</td>
</tr>
<tr>
<td>Clauses with discourse-related gestures</td>
<td>376</td>
</tr>
<tr>
<td>Clauses without gestures</td>
<td>614</td>
</tr>
<tr>
<td>TOTAL</td>
<td>1241</td>
</tr>
</tbody>
</table>

Overall, around half of the clauses were accompanied by co-speech gestures in the corpus, as shown by the frequency of clauses without gestures in the table, that is, 49%. Specifically, discourse-related gestures were the most frequent (30%), followed by representational gestures (14%) and then deictic gestures (7%). This result suggests that more attention needs to be given to discourse-related gestures in future research, both in terms of gesture modeling and empirical gesture studies.

In addition, it is worth noting that the representational gestural rate of basic constructions found in conversations is in stark contrast to the rate found in elicited narratives. Employing the narrative data elicited by ‘Tomato man movies’ (all clips depicting motion events consisting of two protagonists – red Tomato Man and green Triangle man – and a landscape where these protagonists were moving; task: the participants were asked to retell what happened in the clips), Sekine and Kita (2015) reported that the average representational gesture per clause was almost 100%. Their result means that, when asked to retell motion event stories in the video clips, participants made on average one representational gesture together with every clause. A comparison of Sekine and Kita’s study and the present study in this aspect indicates that the difference in representational gesture rates in interviews (14%) and in narratives (100%) is surprisingly huge.33 The huge difference in gestural rate in turn might lead to differences in other aspects of gesture. While existing research on gesture and grammar is mainly concerned with narratives, the present finding suggests that researchers might need to consider conversational data as well to

---

32 Note that unlike Sekine and Kita’s study, the present study does not take deictic gestures as part of representational gestures. However, even though deictic gestures are included in representational gestures in this study, the difference between ‘representational’ gesture rates in their study and this study is still obviously huge.
obtain more results about gestural use beyond that found in elicited narrative descriptions of motion events.

5.3.3 Gesture rates of various categories accompanying high-transitive, intransitive, low-transitive, and copular constructions in speech

The present section considers the use of gestures of various categories co-occurring with each type of transitivity in speech in the corpus. The results are displayed in Table 5.4.

Table 5.4 Representational (abbreviated as Rep.), deictic, and discourse-related gestures (abbreviated as discourse) accompanying the four constructions in speech

<table>
<thead>
<tr>
<th>Type of construction</th>
<th>Rep.</th>
<th>Deictic</th>
<th>Discourse</th>
<th>No gesture</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td>High TRAN</td>
<td>48</td>
<td>19</td>
<td>16</td>
<td>6</td>
<td>72</td>
</tr>
<tr>
<td>INTR</td>
<td>56</td>
<td>25</td>
<td>21</td>
<td>9</td>
<td>56</td>
</tr>
<tr>
<td>Low TRAN</td>
<td>28</td>
<td>10</td>
<td>12</td>
<td>5</td>
<td>96</td>
</tr>
<tr>
<td>COPULAR</td>
<td>37</td>
<td>8</td>
<td>33</td>
<td>7</td>
<td>152</td>
</tr>
</tbody>
</table>

Concerning representational gestures, those accompanying high-transitive constructions (19%) were around two or more times more frequent than those accompanying low-transitive constructions (10%) and copular constructions (8%). The differences between them are statistically significant (high-transitive & low-transitive: p<0.01, \( \chi^2=7.84 \), df=1, phi=0.12; high-transitive & copular: p<0.01, \( \chi^2=20.31 \), df=1, phi=0.17). Similarly, representational gestures co-ocurred with intransitive constructions much more frequently (25%) than with low-transitive (10%) or copular constructions (8%). The differences between them are also statistically significant (intransitive & low-transitive: p<0.01, \( \chi^2=18.58 \), df=1, phi=0.19; intransitive & copular: p<0.01, \( \chi^2=39.4 \), df=1, phi=0.24). However, the representational gesture frequency of intransitive constructions does not significantly differ from that of high-transitive constructions (p>0.05, \( \chi^2=2.50 \), df=1). A similar pattern held in terms of gestural rates of low-transitive constructions and copular constructions (p>0.05, \( \chi^2=1.48 \), df=1). Taken together, there is a general pattern that speakers are more likely to make representational gestures together with high-transitive and intransitive clauses than they are with low-transitive and
copular clauses, but the frequencies of representational gestures accompanying the high-transitive construction and the intransitive construction do not differ significantly. Likewise, the frequencies of representational gestures co-occurring with low-transitive constructions and copular constructions are also similar.

The above results of sharp contrasts of representational gestural rates between intransitive/high-transitive clauses and low-transitive/copular clauses might be due to the different degrees of dynamicity afforded by the two sets of constructions: high-transitive and intransitive clauses prototypically afford a dynamic event construal, whereas low-transitive and copular clauses prototypically afford a static one. In other words, the frequency of the accompanying gestures produced seems to relate to the dynamicity in event construal afforded by a construction. This result provides some empirical evidence in support of the following prediction in the framework of Gestures as Simulated Action (Hostetter & Alibali 2008): the likelihood of gesture production increases from the simulation of perception to the simulation of action.

Furthermore, as noted in Chapter 4 with respect to the means of event construal associated with the four constructions, the event construal afforded by high-transitive constructions is claimed to concern a longer and accordingly more complex action chain (that is, involving an external causation) than that afforded by intransitive constructions is. The same is true of the event construal afforded by the transitive construction compared with that by the copular construction. The similar representational gestural rates of the two sets of constructions (that is, high-transitive vs. intransitive constructions, low-transitive vs. copular constructions) found above seem to suggest that the frequency of gestures produced tends not to be sensitive to the external causation or complexity of event construal (in terms of the length of an action chain). To put it differently, it appears that the external causation or complexity of speakers’ event construal tends not to be reflected in the frequency of the accompanying gestures used. This result is at odds with the Information Packaging Hypothesis, which predicts that patterns that are more difficult to conceptualize tend to elicit more gestures (Kita 2000). Instead, it might be the case that the relation of dynamicity of construal to gestural frequency, as discussed above, could sometimes bypass the relation of complexity of construal to gestural frequency.

In addition, low-transitive constructions are claimed to afford a mentally or subjectively dynamic interaction between two participants, whereas copular constructions are not. However, the similar representational gestural rates of low-transitive and copular constructions as shown above indicate that the frequency of representational gestures produced seems not to correlate with the mental
interaction involved in the construal of events. In other words, the mental interaction involved in the construal of events may not be reflected in the frequency of the accompanying representational gestures produced.

With regard to deictic gestures accompanying various types of constructions in speech, no significant correlation was found (p>0.05, $\chi^2=4.86$, df=3). The similarity in the number of deictic gestures accompanying all types of constructions shows that gestures of this type tend not to reflect the different ways in which speakers conceptualize these constructions. This furthermore suggests a different cognitive origin for representational gestures and deictic gestures. Thus, representational gestures and deictic gestures can best be identified separately in future gesture research, which is one of the motivations for the coming chapters, which simply focus on representational gestures.

Turning now to discourse-related gestures, these gestures seem to be slightly more likely to accompany low-transitive constructions (35%) and copular constructions (31%) than high-transitive constructions (28%) and intransitive constructions (25%), but only one difference was found to be statistically significant: the difference between the frequency of gestures occurring with low-transitive constructions (35%) and that with intransitive constructions (25%) (p<0.05, $\chi^2=6.31$, df=1, phi=0.11). A comparatively more frequent use of discourse-related gestures together with low-transitive constructions might be motivated by pragmatic factors. For instance, low-transitive clauses appear to involve relatively more clauses consisting of mental space builders, such as I think/I know/hope + that clause, which were found by Cienki (2009) to be relatively frequently accompanied by gestures. Nevertheless, the frequency of discourse-related gestures produced seems not to be linked with the ways of conceptualizing events, neither with respect to the dynamicity, the complexity of the construal of the four constructions (that is, external causation) nor with respect to the kind of mental interaction involved (that is, mental contact claimed to exist between two participants in low-transitive clauses, such as the mental contact between I and the bird afforded by the utterance I saw a bird in the sky).

Taken together, two major findings emerge from this part. One is that the frequency of representational gestures produced correlates in different ways with the four constructions, although the frequency of deictic gestures or discourse-related gestures does not. This result could be a consequence of the

---

34 Gesture models usually collapse across these gestures and refer to them as referential gestures, such as in the model of Gestures as Simulated Action.

35 Note that Cienki (2009) found that head movements were more frequent than hand movements to occur with expressions of mental space builders.
different means of event construal afforded by the four constructions. The other finding is that speakers tend to utter more clauses with lower transitivity (low-transitive and copular clauses) without representational gestures, whereas they tend to produce many more representational gestures accompanying clauses with higher transitivity (high-transitive and intransitive clauses). Simply put, transitivity is low in conversational speech but high in terms of its expression in gesture.

5.4 Modes of Representation of representational gestures with respect to high-transitive, intransitive, low-transitive, and copular constructions

In what follows, I will present how speakers gestured when they used the four constructions in speech and then discuss to what extent these gestures relate to the way in which a speaker may be conceptualizing each construction. An overview of the results is shown in Table 5.5.

Table 5.5: Various gestural Modes of Representation (MoR) with respect to various types of transitivity in speech

<table>
<thead>
<tr>
<th>Type of construction</th>
<th>Acting with object</th>
<th>Acting only</th>
<th>Tracing</th>
<th>Molding</th>
<th>Embodying</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>High TRAN</td>
<td>38</td>
<td>41</td>
<td>4</td>
<td>4</td>
<td>26</td>
<td>28</td>
</tr>
<tr>
<td>INTR</td>
<td>13</td>
<td>13</td>
<td>25</td>
<td>26</td>
<td>48</td>
<td>50</td>
</tr>
<tr>
<td>Low TRAN</td>
<td>11</td>
<td>19</td>
<td>3</td>
<td>5</td>
<td>9</td>
<td>16</td>
</tr>
<tr>
<td>COPULAR</td>
<td>9</td>
<td>13</td>
<td>2</td>
<td>3</td>
<td>12</td>
<td>18</td>
</tr>
</tbody>
</table>

From this table we can see frequent gestural modes with respect to the four constructions in speech, in particular Acting-with-object gestures occurring with high-transitive constructions, Acting-only and Tracing gestures with intransitive constructions, and Molding gestures (and Embodying gestures) with low-transitive and copular constructions. The frequent gesture-speech co-expressions and the relevant statistical results are provided below.

5.4.1 Gestures of various modes co-occurring with high-transitive constructions

Chi-square tests were carried out for all the above gestural Modes of Representation accompanying high-transitive clauses and the other types of clauses in the corpus.
Only one positive correlation was obtained: Acting-with-object gestures were found to significantly correlate with the high-transitive clauses \((p<0.01, \chi^2=23.32, \text{df}=1, \phi=0.27)\). An example of this gestural mode occurring with a high-transitive clause is as follows. When a speaker described an activity whereby he put a battery into a microphone, he said, \textit{put it in (the microphone)}, and at the same time he made the gesture shown in Figure 5.2. 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.

![Figure 5.2 Gesture with put it (battery) in (the microphone)](image)

Based upon the Gestures-as-Simulated-Action framework (Hostetter & Alibali 2008), 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 relates to the hypothesized means by which a speaker may conceptualize a high-transitive construction. Langacker (1990, 2008a) 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. On the one hand, transmission of force implies ‘relative control’ of the former participant over the latter one (Croft 2012: 230); 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.

### 5.4.2 Gestures of various modes co-occurring with intransitive constructions

Chi-square tests showed that intransitive clauses tend to correlate with the use of Tracing \((p<0.01, \chi^2=23.79, \text{df}=1, \phi=0.27)\) and Acting-only gestures \((p<0.01, \chi^2=30.64, \text{df}=1, \phi=0.27)\).
Tracing gestures (50%) were more frequent than Acting-only ones (26%).

Tracing gestures occurring with intransitive clauses usually depicted a trajectory movement of an entity. For example, when a speaker said *The microphone comes, flying out of the sky*, he was tracing a trajectory in space, and his finger was embodying a microphone, as shown in Figure 5.3. In this gesture, neither the manner of the microphone’s motion (that is, the flying motion) nor the external force that caused the movement of the microphone (that is, someone tossing it to the speaker in this case) was foregrounded. Instead, only the microphone’s change of location was depicted in the gesture.

![Figure 5.3 Gesture occurring with (The microphone) flying out of the sky](image)

Acting-only gestures depicted the manner of a motion activity. For example, while a speaker uttered *I just kind of, crawl on the bed*, she mimed the activity of crawling, as shown in Figure 5.4. She moved her whole upper body as if to mime the crawling activity, which involved hands/arms, head, and eyes. What was foregrounded in this gesture was mainly the manner of the agentive motion activity.

![Figure 5.4 Gesture with I just kind of, crawl on the bed](image)

Gestures of both modes above, depicting either the trajectory or the manner of a motion, expressed non-object-oriented manual activities. That is to say, they only expressed entities’ movements, rather than any external causation. These appear to
correlate with 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 5.5 above, Acting-only gestures were much less preferred by speakers than Tracing gestures to accompany intransitive clauses. The question then arises as to whether this might relate to the sub-types of intransitive clauses (associated with different means of event construal). As noted in Chapter 4, event construal afforded by intransitive clauses can be classified into two sub-kinds: a) absolute construal of events, in which speakers only focus on the change of state/location of a participant (usually in terms of non-agentive motions, such as *the ring dropped on the floor*); b) force-dynamic construal (that is, internal causation/agentive) of events, in which speakers focus on an internal cause of a motion activity (besides the change of state or location it brings about), such as *she’s dancing* and *she went away*. It would be of interest to explore whether the (different) uses of Acting-only and Tracing gestures relate to the above sub-types of intransitive constructions in speech. I will now examine this question in more detail.

a) **Acting-only and Tracing gestures occurring with agentive and non-agentive intransitive constructions**

In order to determine whether or not the use of Acting-only and Tracing gestures relates to the type of intransitive clauses, I classified intransitive clauses accompanied by these gestures (Acting-only and Tracing gestures) into two types: agentive ones (e.g., *she danced all night long*) and non-agentive ones (e.g., *she fell deeply in love*). Their relations to the accompanying gestures are shown in Table 5.6.

<table>
<thead>
<tr>
<th>Type of intransitive clauses</th>
<th>Acting only</th>
<th>Tracing</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td>Agentive</td>
<td>21</td>
<td>38</td>
<td>34</td>
</tr>
<tr>
<td>Non-agentive</td>
<td>5</td>
<td>26</td>
<td>14</td>
</tr>
</tbody>
</table>

Table 5.6 shows that both types of intransitive clauses were more frequently accompanied by Tracing gestures rather than Acting-only gestures, that is, 62% (vs. 38%) and 74% (vs. 26%), respectively. There is no significant difference as for the uses of Tracing or Acting-only gestures with respect to the two types of intransitive
clauses ($p>0.05$, $\chi^2=0.43$, df=1). This result suggests that speakers tend not to
differentiate between two types of intransitive clauses (affording either
agentive/force-dynamic or non-agentive/absolute event construal) via the use of
gestural Modes of Representation (that is, Acting-only and Tracing gestures).

Given that the use of Acting-only and Tracing gestures tends not to relate to the
agentivity of intransitive constructions, it therefore remains a puzzle as to why
speakers preferred to use Tracing gestures more often versus Acting-only gestures.
The coming sub-section will involve a tentative exploration of this question, in the
hope of gaining a more comprehensive understanding of gestural use with respect to
intransitive constructions in spoken language.

b) A further, tentative exploration of the use of Acting-only and Tracing gestures

Previous studies suggest that the language one speaks influences the gestural
representation when depicting motion events (Kita & Özyürek 2003; McNeill 1997,
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. This
section aims to 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 coded 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) (Levin 1993), 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”
Basic clausal constructions (transitivity) and gesture

(abbreviated as NoMo.), such as activities of appearance/disappearance (e.g., appearing), posture activities (e.g., sitting), and body-internal activities (e.g. waving). Table 5.7 displays the results of these types of motion components in relation to the use of Acting-only and Tracing gestures.

Table 5.7 Acting-only and Tracing gestures occurring with intransitive clauses with “manner only”, “path only”, and the “combination” of the two, as well as those without any motion events (NoMo.)

<table>
<thead>
<tr>
<th>Type of intransitive clauses</th>
<th>Acting only</th>
<th>Tracing</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td>Manner only</td>
<td>7</td>
<td>64</td>
<td>4</td>
</tr>
<tr>
<td>Path only</td>
<td>6</td>
<td>17</td>
<td>29</td>
</tr>
<tr>
<td>Combination</td>
<td>2</td>
<td>15</td>
<td>11</td>
</tr>
<tr>
<td>NoMo.</td>
<td>9</td>
<td>69</td>
<td>4</td>
</tr>
<tr>
<td>TOTAL</td>
<td>24</td>
<td>33</td>
<td>48</td>
</tr>
</tbody>
</table>

From the table, we can see 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.

Table 5.7 also shows 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%. These differences are statistically significant (“manner only” & “path only”: p<0.01, χ²=6.78, df=1, phi=0.38; “manner only” & “combination”: p<0.05, χ²=5.49, df=1, phi=0.48; “path only” & “NoMo.”: 0.05 < p < 0.1, χ²=3.96, df=1, phi=0.34).
p<0.01, $\chi^2=9.67$, df=1, phi=0.45; “combination” & “NoMo.”: p<0.05, $\chi^2=5.67$, df=1, phi=0.47). 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.

5.4.3 Gestures of various modes co-occurring with low-transitive constructions
Chi-square tests for all the gestural modes accompanying low-transitive clauses and other types of clauses indicated that 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). One example of the use of Molding gestures is as follows. While a speaker said The carpet has a Mickey Mouse in it, she was molding the form of a Mickey Mouse image in gesture, as shown in Figure 5.5. What was profiled in the gesture was only an entity – “Mickey Mouse” – rather than any dynamic activities related to it.

Figure 5.5 Gesture with The carpet has a Mickey Mouse
Molding gestures, which usually simply depicted entities, seem to be associated with one means of event construal afforded by low-transitive constructions but not

36 Note that these results remain speculative, given that the number of intransitive clauses expressing the “manner only”, the “combination” of manner and path and those do not express the motion events (“NoMo.”) is not big enough in the data.
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 5.6. 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 (Langacker 1990, 2008a). In the gesture produced, the speaker moved and held her fists, as if she were grasping 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 grasping movements.

![Figure 5.6 Gesture with I want this](image)

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 (1990, 2008a).

5.4.4 Gestures of various modes co-occurring with copular constructions
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. Gestures of these two modes usually just profiled entities rather than the related dynamic activities. An example of a Molding gesture is the following. When a speaker said *He’s a domestic cat*, she made the gesture shown in Figure 5.7. She held her two hands apart, palms facing each other with tensed, slightly curved and spread fingers, as if she were touching the cat. Note that the speaker only referred to the static entity here, rather than to any activities related to the cat, although in theory she could have made gestures in that way instead (such as imitating the cat walking, for example).

![Figure 5.7 Gesture with He’s a domestic cat](image)

As noted in Chapter 4, 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).
To summarize, all the above analyses show that the modes of gestural representation preferably correlate with 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.

5.5 Form parameters of discourse-related gestures in relation to high-transitive, intransitive, low-transitive, and copular constructions

This section considers how speakers produced discourse-related gestures when they used the four constructions in speech. It aims to determine whether the use of discourse-related gestures is associated with the type of construction in speech (that is associated with various means of event construal, mainly in terms of complexity/external causation, dynamicity, and mental interaction, as just discussed), which can further suggest whether representational gestures and discourse-related gestures possibly share a common cognitive origin.

The results for the tension or lack of tension in discourse-related gestures co-occurring with the four basic constructions are displayed in Table 5.8.

Table 5.8 Discourse-related gestures with and without tension co-occurring with the four constructions

<table>
<thead>
<tr>
<th>Type of construction</th>
<th>Gesture with tension</th>
<th></th>
<th>Gesture without tension</th>
<th></th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td>High TRAN</td>
<td>27</td>
<td>37</td>
<td>45</td>
<td>63</td>
<td>72</td>
</tr>
<tr>
<td>INTR</td>
<td>23</td>
<td>41</td>
<td>33</td>
<td>59</td>
<td>56</td>
</tr>
<tr>
<td>Low TRAN</td>
<td>31</td>
<td>32</td>
<td>65</td>
<td>68</td>
<td>96</td>
</tr>
<tr>
<td>COPULAR</td>
<td>53</td>
<td>35</td>
<td>99</td>
<td>65</td>
<td>152</td>
</tr>
</tbody>
</table>

We see that, overall, speakers tended to produce discourse-related gestures without tension in a majority of cases – around 60%, regardless of the type of construction the gestures accompanied. A statistical analysis of these gestures accompanying the four constructions showed that there is no significant difference
among the gestures ($p>0.05$, $\chi^2=1.34$, df=3), meaning that this form parameter of discourse-related gestures tends not to correlate with the type of constructions in the dataset.

Next, the movement sizes (including finger, wrist, forearm, and arm) of discourse-related gestures co-occurring with these constructions were examined. Given the rare use of finger and arm movements in making discourse-related gestures, they were collapsed into the categories wrist and forearm respectively. The results are shown in Table 5.9.

Table 5.9 Discourse-related gestures accompanying the four constructions with different gestural sizes

<table>
<thead>
<tr>
<th>Type of construction</th>
<th>Wrist gesture</th>
<th>Forearm gesture</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td>High TRAN</td>
<td>22</td>
<td>31</td>
<td>50</td>
</tr>
<tr>
<td>INTR</td>
<td>14</td>
<td>25</td>
<td>42</td>
</tr>
<tr>
<td>Low TRAN</td>
<td>29</td>
<td>30</td>
<td>67</td>
</tr>
<tr>
<td>COPULAR</td>
<td>41</td>
<td>27</td>
<td>111</td>
</tr>
</tbody>
</table>

We see that, overall in the corpus, speakers tended to make discourse-related gestures bending the forearm more often than they did bending the wrist. A significant correlation between the type of gesture size and the type of construction in speech was actually not found ($p>0.05$, $\chi^2=0.78$, df=3).

The above results indicate that two form parameters of discourse-related gestures – gestural size and tension, which were suspected to be most promising in terms of relating to the type of construction – were not found to correlate with the type of construction in speech. This result suggests that the forms of discourse-related gestures seem less likely to relate to the means of construal of these constructions, at least not in terms of the above movement qualities. To put it differently, the activation which evokes the use of utterances with the different types of transitivity tends not to motivate the forms of discourse-related gestures.

These results complement those in Masson-Carro et al. (2016). Masson-Carro and colleagues found that the representational gestural rate is influenced by the degree of action affordance of objects, whereas the discourse-related gestural rate is not. This suggests that different types of gesture may bear different cognitive origins. Yet that study only investigated the rate of discourse-related gestures, without
considering their forms – another dimension which might be primed by the ‘general activation’. By including both the production rate (as shown in Section 5.3) and the forms of discourse-related gestures (as shown in the current section), the current study found that neither the rates nor the forms of discourse-related gestures were related to the type of transitivity. All of these findings suggest that representational gestures and discourse-related gestures may have different cognitive origins, thus going against the ‘general activation’ account.

5.6 Summary and conclusions
The current chapter has investigated a) the use of four basic constructions in speech and the use of gestures accompanying these four constructions, and b) the question to what extent the gestures used correlated with the ways of conceptualizing the constructions (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 speakers preferred to utter clauses with lower transitivity (that is, low-transitive and copular clauses), whereas they preferably made representational gestures accompanying clauses with higher transitivity (that is, high-transitive and intransitive clauses). This finding of lower transitivity in speech and higher transitivity in representational gestures indicates a mismatch between gesture use and speech use in interaction along the dimension of transitivity. This also suggests that the discussion about the extent of multimodality of grammar should not be taken too far; that is to say, a radical multimodal approach to grammar would not be an appropriate direction to pursue. Instead, this finding, together with another finding that different constructions involve significantly different gestural rates, suggests that a “variably multimodal” approach, in terms of a variably greater or lesser degree of multimodality according to various factors, seems more plausible (see Cienki 2015).

Analyses of frequencies and forms of representational, deictic, and discourse-related gestures accompanying the four constructions showed that what correlated with the ways in which speakers conceptualize the constructions was restricted to representational gestures. Major relations between gesture and speech found in this study are summarized in Table 5.10. This suggests that the three types of gestures may have different cognitive origins, thereby going against the ‘general activation’ account. Thus, the argument for multimodality of a certain grammatical construction, which is made in research which collapses gestures of various categories together, might not be equally tenable for all types of gestures.

Furthermore, concerning representational gestures, their frequency and forms were found to correlate with the construal of the constructions these gestures
accompanied. Specifically, the dynamicity involved in constructions seems to be reflected in both the rate of representational gestures and the preferred depicting modes, whereas the external causation involved in these constructions tends to be only reflected in the preferred depicting modes of the representational gestures, but not in the frequency of gestures produced. However, compared with the above two dimensions, the dimension of “mental contact” (involved in the low-transitive construction) appears less likely to be reflected either in the gestural frequency or in the preferred depicting modes of the representational gestures. All these findings indicate that representational gestures (including their rates and/or forms) largely correlate with different means of event construal in different ways. The correlation between gesture and the construal of events found in this study corroborates the view that both spoken language and gesture are expressions of forms of construal. The current findings are furthermore informative about the ways in which gesture could express different kinds of event construal.

Table 5.10 A summary of correlations found between ways of conceptualizing transitivity and behaviors of gestures in this chapter (“+” represents there is a significant correlation; “−” represents there is not).

<table>
<thead>
<tr>
<th>Types of event construal</th>
<th>Construction in speech</th>
<th>Co-speech gesture</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>High TRAN</td>
<td>INTR</td>
</tr>
<tr>
<td>Dynamicity</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>External causation</td>
<td>+</td>
<td>−</td>
</tr>
<tr>
<td>Mental contact</td>
<td>+</td>
<td>−</td>
</tr>
</tbody>
</table>

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 (1990, 2008a). Whereas the proposed difference in conceptualizing transitive and intransitive constructions (that is, with or without a profile of an external causation, or the length of the Action Chain, or simply the complexity of construal) has gained support in previous imaging studies (Thompson et al. 1997), 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\textsuperscript{37} which is related to this means of construal 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.

Besides the above correlations between gesture and the means of event construal afforded by the four constructions, the use of Acting-only and Tracing gestures occurring with intransitive clauses was further investigated in relation to the sub-types of event construal afforded by the intransitive construction (that is, +/-agentive) and to the manner modulation, as suggested by previous studies. It was found that the use of these gestural modes (that is, the Acting-only and Tracing modes) seems not to relate to the sub-type of intransitive clause, and that manner modulation appears not to be an appropriate account for the more frequent use of Tracing gestures occurring with intransitive clauses. Instead, the findings indicated that intransitive clauses expressing the path (including the “path only” and the “combination” of path and manner) tend to be accompanied by Tracing gestures, whereas intransitive clauses expressing “manner only” tend to be accompanied by the Acting-only gestures. In addition, a further, important finding was that the more frequent use of Tracing gestures rather than the Acting-only gestures occurring with intransitive motion events in English is simply due to the more frequent use of intransitive clauses expressing “path only” (together with the “combination”) versus intransitive clauses expressing “manner only”. This result is in conflict with the Manner Modulation Hypothesis, as proposed in previous studies to account for the use of Tracing gestures versus Acting-only gestures in English. The above results offer a more comprehensive understanding of gestural use in relation to intransitive clauses.

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 correlate with 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., \textit{she broke the window} vs. \textit{the wind broke the window}), intentionality of the subject (e.g., \textit{she baked a cake} vs. \textit{she dropped the cake on the floor}), affectedness of the object (e.g., \textit{she put it on the table} vs. \textit{she’s holding it in her hand}), and others (a list of parameters which have been used to categorize transitivity can be found in Chapter 4). It would be worthwhile to examine to what extent gesture correlates with these sub-types of...

\textsuperscript{37}Note that this gestural representation (Acting-with-object gesture), which can reflect the abstract conceptualization of low transitive constructions, is less preferred than the other another type of gesture is (Molding gesture) in the corpus.
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. Furthermore, the relation between discourse-related gestures and the construal of transitivity (or other grammatical categories) can be explored in a more comprehensive, objective way. For instance, a greater variety of manual movement parameters (such as movement types, movement qualities, and others) could be examined with motion-capture data, which could be more objective in measuring these facets in manual movements.