
6.1 Introduction

Polysemy refers to “the association of two or more related senses with a single linguistic form” (Taylor 1995: 99). As a ubiquitous phenomenon in natural language and “the pivot of semantic analysis” (Ullmann 1951: 117), polysemy has gained wide attention in Cognitive Linguistics. The prototype approach to word meaning in Cognitive Linguistics proposed that the senses of a polysemous word center around a prototype and are systematically interrelated by a family resemblance structure, and that they are organized as a schematic network via links of schematization and extension (e.g., Lakoff 1987; Lakoff & Johnson 1980; Langacker 1987, 1991a, 1991b). Many studies found that polysemy involves the extension of a central meaning to other non-central meanings by means of basic cognitive mechanisms such as metaphor, metonymy, image schema, proposition, generalization, and specialization etc. (e.g., Brugman & Lakoff 1988; Johnson 1987; Lakoff 1987). According to Cognitive Linguistics, meaning is equated with conceptualization, and grammar is meaningful by nature (Langacker 1987, 1991a, 1991b, 2008a). Like lexical items, grammatical items are frequently polysemous. However, their meanings are often abstract and reside in our conceptualization (Langacker 2008b: 16). As the only explicitly marked aspectual form in English, the progressive construction (be V-ing) has developed as a polysemous construction, resulting in a rather complex semantic network (Bybee et al. 1994; Dahl 1985; De Wit & Brisard 2014; Langacker 1987, 2001, 2008a).

The use of the English progressive construction has temporal and modal motivations. While the aspecto-temporal uses of the progressive construction have the
semantic potentials of expressing current ongoingness, habituality, iteration, and temporary state, its modal (i.e., non-aspecto-temporal, subjective) uses give rise to the futurate reading or the epistemic readings of surprise, tentativeness and intensification (see section 2.2.1 in Chapter 2 for details). As the prototypical use of the progressive construction, the progressive aspect is a grammatical device for “zooming in” and taking an “internal perspective” on an event, as if the conceptualizer were watching it unfold sequentially rather than viewing it holistically as a whole (Comrie 1976; Langacker 1991; Langacker 2001; Smith 1997). Grammatical aspect categories not only depict grammatical features but also mirror conceptual structures, and hence have psychological reality. Psychological studies have shown that the English progressive aspect plays a central role in mental simulation. It facilitates language comprehenders’ mental simulation of actions and locations, producing a large Action-sentence Compatibility Effect and Location-sentence Compatibility Effect (Bergen & Wheeler 2010; Glenberg & Kaschak 2002; Liu & Bergen 2016). The English progressive aspect also activates language comprehenders to conceptualize more action in a given time (Huette et al. 2012; Madden & Zwaan 2003; Matlock 2010, 2011), and on the part of speakers, its use correlates with the production of longer and more complex gestures (Duncan 2002; Parrill et al. 2013). The previous studies only center on the “continuous ongoingness” of progressive aspect uses in motion events. It is unclear whether and how the other senses of the progressive construction affect speakers’ mental simulation in natural conversations.

Speech and co-speech gestures have been claimed to be an integrated system (e.g., Kendon 2004; McNeill 1992, 2005). If so, one might expect that the gestures accompanying utterances with different meanings of the progressive construction should be distinct from each other. The GSA framework proposes that gestures derive from internal motor or visual simulation of action (Hostetler & Alibali 2008). Representational gestures – representing concrete or abstract objects, actions, and ideas – are regarded as “outward manifestations of the action components involved in mental simulations of spatial and motor events” (Hostetler 2014: 1469). The meaning of the speaker’s message is “externally embodied in representational gestures” (Marghetis &
Bergen 2014: 2000). Therefore, this chapter is interested in investigating to what extent and how the different meanings of the progressive construction are externally embodied in gestures. The specific questions are as follows.

1) To what degree is the polysemous progressive construction in modern spoken English multimodal?

2) What are the prototypical gestures accompanying the different meanings of the progressive construction?

3) What are the gestural characteristics of iconic gestures accompanying the English progressive aspect?

Answering these questions will indicate whether the different functions of the progressive aspect in English have different multimodal manifestations or not. If so, gesture can provide insight into how the mental simulations behind the different functions differ – an important point if we take the cognitive linguistic approach to meaning as conceptualization.

6.2 Methods

6.2.1 Corpus

The data used in this research come from “The Ellen DeGeneres Show”. I had access to the data via the UCLA Library Broadcast NewsScape via the Distributed Little Red Hen Database, co-directed by Francis Steen and Mark Turner. In order to extract the utterances with a “be V-ing” construction only from The Ellen DeGeneres Show, I created a subcorpus–ellen degeneres show– in CQPweb, a web-based corpus analysis system combining power, flexibility and usability (Hardie 2012). Then I carried out a standard query in the form of “{be/V} (_RB)? _VBG”, developed by Peter Uhrig. Since this cannot extract out the abbreviated forms like “’s/’re V-ing”, I obtained 77,545 matches in 1,638 different texts. Then I exported them into Excel and manually eliminated the instances which have no progressive “be V-ing” constructions as in (1a), the equative constructions containing main verb BE and gerund as in (1b), and the -ing
forms used as predicative adjectives as in (1c). I have also left out all the be going to constructions as in (1d), because it has been grammaticalized into the expression of future tense.

(1) a. His new movie is “The Hateful Eight”. (Ellen Show #913)
   b. And one of my favorite things is waiting for the show night to see what flowers Chris will send her. (Ellen Show #138).
   c. The end result is something that is very captivating (Ellen Show # 294)
   d. While you look at how cute it is, I am going to take a nap. (Ellen Show # 1)

In addition, I also eliminated those constructions with progressive forms from commercial advertisements, song scripts, the breaking news report, the game unit and when the speakers were invisible, because they all involved different kinds of demonstrations, showing other images, or real actions without equal opportunities for gestural expression.

As gesture coding is laborious and time-consuming work, I just examined the first 200 progressive utterances made by the guests and the hostess respectively in the interview part, leaving out all the data from monologue and game parts. Thus, the corpus consists of 400 progressive utterances in total. As Ellen has a different role in the interaction from the guests, participating in each episode, I separated her data from that of guests for analyses to avoid individual bias. The 400 progressive utterances are from the interview between Ellen and guests of 35 episodes featured from November 10, 2015 to January 19, 2016. Each interview lasts for about 20 minutes, so the whole corpus consists of about 700 minutes of material.

6.2.2 Speech coding

In terms of the meanings of the progressive construction (see section 2.1.2.1 in Chapter 2 for details), each utterance was categorized into “continuous ongoingness”, “habituality”, “repetition”, “temporary state”, “futurity” or “epistemicity”. Two coders
annotated all 400 utterances. The agreement for the meaning annotation was 92% (N = 400).

6.2.3 Gesture coding

Each progressive utterance was coded for whether it occurred with gesture or without gesture. Based on the previous studies on gesture types (Kendon 2004; McNeill 1992, 2005), each gesture was annotated as deictic, concrete iconic, metaphoric iconic or pragmatic. Twenty percent of the progressive utterances with gestures (Total N = 239) were coded by the second rater. Then the iconic gestures (concrete and metaphoric) were flagged out, and further annotated (N = 60). In terms of the semantic elements it was iconic to, each iconic or metaphoric gesture was coded for iconicity to entity, action, path of motion, or abstract process involved in an event. There were five utterances that were accompanied by multiple gestures; I analyzed only the gesture that was synchronous with the progressive construction. All of the iconic gestures were coded by two raters according to the coding manual. Inter-rater reliability of this judgment across two independent observers was 93% for gesture type, and 97% for the semantic elements the iconic gestures are iconic to.

6.3 Results

I report the results of the guests as a community and the hostess ELLEN as an individual in separate ways as follows. To find out the potential multimodality of different meanings of the progressive construction, I compared the frequency distributions of the gestures accompanying them. Then to investigate the role the gestures play, I compared the frequency distributions of their functions.

6.3.1 Frequency of actual usage and the multimodality of the progressive construction in spoken English

Table 6.1 shows the distribution of gestures in the different usage types of the
progressive construction in our dataset. Overall, the aspecto-temporal use dominates in the conversation context, while the non-aspectual use is relatively not frequent (79.5% vs. 20.5%). More than one third of the utterances with the progressive construction triggered the “continuous ongoingness” interpretation (37.0%), which is the prototypical meaning of this construction. This result is consistent with previous studies of the progressive construction uses in an oral corpus or a mixed corpus of both spoken and written data (De Wit & Brisard 2014; Rompaey 2014). “Temporary state” use occurred even more frequently than the “habituality” reading (25.5% vs. 17.0%). There was no use of “repetition” with be V-ing, a use which was also found to be marginal in other studies (De Wit & Brisard 2014; Rompaey 2014). The non-aspectual use of be V-ing includes almost twice as many instances of “emotivity” (13.0%) as “futurity” (7.5%).

<table>
<thead>
<tr>
<th>Meanings</th>
<th>With gesture</th>
<th>Without gesture</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td>Ongoingness</td>
<td>51</td>
<td>25.5</td>
<td>23</td>
</tr>
<tr>
<td>Habituality</td>
<td>20</td>
<td>10.0</td>
<td>14</td>
</tr>
<tr>
<td>Repetition</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Temporary State</td>
<td>27</td>
<td>13.5</td>
<td>24</td>
</tr>
<tr>
<td>Futurity</td>
<td>10</td>
<td>5.0</td>
<td>5</td>
</tr>
<tr>
<td>Emotivity</td>
<td>15</td>
<td>7.5</td>
<td>11</td>
</tr>
<tr>
<td>Total</td>
<td>123</td>
<td>61.5</td>
<td>77</td>
</tr>
</tbody>
</table>

Generally speaking, the progressive construction shows much greater potential to be multimodal than mono-modal (61.5% with gesture vs. 38.5% without gesture). No matter what the meaning of the progressive construction is, the utterances accompanied by gestures are more frequent than those without gestures. The multimodality of the utterances denoting “continuous ongoingness” was the highest (25.5%), and that of the “futurity” was the lowest (5.0%). However, the Chi-square result showed no significant differences of gesture use among the progressive construction with different meanings ($\chi^2(4) = 3.729, p > 0.05$) (the repetition category was excluded).

In Table 6.2, you can see that Ellen also made far more aspecto-temporal use than
non-aspectual use of the progressive construction (64% vs. 36%). However, there is almost equal use of the progressive construction with the meaning of “continuous ongoingness” (21.0%), “temporary state” (23.0%), and “emotivity” (21.5%). There are also very few progressive utterances denoting “repetition”. In contrast with the results of the guests (see Table 6.1), Ellen made far greater use of progressive utterances with non-aspectual meanings than the guests did (36% vs. 20.5%). There is less “continuous ongoingness” use of the progressive construction (21.0% vs. 37.0%). “Temporary state” use is equally high between them (23.0% for Ellen and 25.5% for the guests), and “repetition” use is scarce (1.0% and 0, respectively).

Table 6.2: Gesture use for the meanings of the progressive construction by ELLEN

<table>
<thead>
<tr>
<th>Meanings</th>
<th>With gesture</th>
<th></th>
<th>Without gesture</th>
<th></th>
<th>Total</th>
<th></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>Ongoingness</td>
<td>27</td>
<td>13.5</td>
<td>15</td>
<td>7.5</td>
<td>42</td>
<td>21.0</td>
</tr>
<tr>
<td>Habituality</td>
<td>24</td>
<td>12.0</td>
<td>14</td>
<td>7.0</td>
<td>38</td>
<td>17.0</td>
</tr>
<tr>
<td>Repetition</td>
<td>1</td>
<td>0.5</td>
<td>1</td>
<td>0.5</td>
<td>2</td>
<td>1.0</td>
</tr>
<tr>
<td>Temporary State</td>
<td>25</td>
<td>12.5</td>
<td>21</td>
<td>10.5</td>
<td>46</td>
<td>23.0</td>
</tr>
<tr>
<td>Futurity</td>
<td>15</td>
<td>7.5</td>
<td>14</td>
<td>7.0</td>
<td>29</td>
<td>14.5</td>
</tr>
<tr>
<td>Emotivity</td>
<td>24</td>
<td>12.0</td>
<td>19</td>
<td>9.5</td>
<td>43</td>
<td>21.5</td>
</tr>
<tr>
<td>Total</td>
<td>116</td>
<td>58.0</td>
<td>84</td>
<td>42.0</td>
<td>200</td>
<td>100</td>
</tr>
</tbody>
</table>

In terms of gestural tendency, the progressive utterances denoting “continuous ongoingness”, “habituality” and “emotivity” are more likely to be accompanied by gesture, that is, they have greater potential to be multimodal. However, there is almost 50% chance of the progressive utterances with “temporary state” and “futurity” construal being accompanied by gestures. Overall, however, the differences between meanings and gesture rate are not significant ($\chi^2(4) = 1.954$, $p > 0.05$).

To summarize, the progressive construction as a whole has greater potential to be multimodal than monomodal in our data set. “Continuous ongoingness” as the prototypical meaning of the progressive aspect is dominantly used most often and the most likely to be multimodal. The English progressive construction has expanded to encode “temporary state” in modern English conversations, but only half of these cases
are accompanied by gestures. The hostess Ellen used far more progressive constructions
to encode “futurity” and “emotivity” than the guests did, almost half of which were
multimodal representations.

6.3.2 The characteristics of gestures accompanying the progressive construction

The characteristics of gestures accompanying the progressive construction are revealed
through the general functions and specific functions (i.e., semantic elements) in the
following subsections.

6.3.2.1 Frequencies of gestures accompanying the progressive construction with
different meanings

Table 6.3 presents the distribution of the four gesture types in the five kinds of meanings
denoted by the progressive construction. Overall, the pragmatic gesture accompanying
the progressive utterances dominated in the corpus (44.7%), which is twice as much
with the deictic gesture (21.1%). There are almost equal amounts of concrete iconic
gesture (17.9%) and metaphoric iconic gesture (16.3%), but they are both far less
frequent than pragmatic gesture. Utterances with different meanings of the progressive
construction are most frequently accompanied by pragmatic gestures. However, there
are almost six times and three times more concrete iconic gestures accompanying the
utterances with “continuous ongoingness” construal than with “temporary state” and
with “futurity” respectively (14.6% vs. 2.4% vs. 4.5%). Progressive utterances with the
construal of “temporary state” are more often accompanied by metaphoric iconic
gestures than those of “current ongoingness”, “emotivity” and “habituality”. The
deictic gestures are relatively evenly distributed among the progressive utterances with
different meanings. There are no concrete iconic gestures co-occurring with the
progressive utterances denoting “habituality” and “emotivity”. There is no metaphoric
iconic gesture co-occurring with the “futurity” interpretation of the progressive
utterances. This result suggests that the non-aspectual use of the progressive
construction (i.e., “futurity” and “emotivity”) is not likely to be accompanied by iconic
gestures.

Table 6.3: Gesture types for the meanings of the progressive construction by GUESTS

<table>
<thead>
<tr>
<th>Meanings</th>
<th>Deictic</th>
<th>Iconic</th>
<th>Metaphoric</th>
<th>Pragmatic</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>Ongoingness</td>
<td>5</td>
<td>4.1</td>
<td>18</td>
<td>14.6</td>
<td>7</td>
</tr>
<tr>
<td>Habituality</td>
<td>4</td>
<td>3.3</td>
<td>0</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Tem. state</td>
<td>7</td>
<td>5.7</td>
<td>3</td>
<td>2.4</td>
<td>8</td>
</tr>
<tr>
<td>Futurity</td>
<td>4</td>
<td>3.3</td>
<td>1</td>
<td>4.5</td>
<td>0</td>
</tr>
<tr>
<td>Emotivity</td>
<td>6</td>
<td>4.9</td>
<td>0</td>
<td>1</td>
<td>5.0</td>
</tr>
<tr>
<td>Total</td>
<td>26</td>
<td>21.1</td>
<td>22</td>
<td>17.9</td>
<td>20</td>
</tr>
</tbody>
</table>

* Notes: Tem is short for Temporary.

Table 6.4: Gesture types for the meanings of the progressive construction by ELLEN

<table>
<thead>
<tr>
<th>Meanings</th>
<th>Deictic</th>
<th>Iconic</th>
<th>Metaphoric</th>
<th>Pragmatic</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>Ongoingness</td>
<td>12</td>
<td>10.3</td>
<td>7</td>
<td>6.0</td>
<td>1</td>
</tr>
<tr>
<td>Habituality</td>
<td>8</td>
<td>6.9</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Repetition</td>
<td>1</td>
<td>0.9</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Tem. state</td>
<td>8</td>
<td>6.9</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Futurity</td>
<td>8</td>
<td>6.9</td>
<td>1</td>
<td>0.9</td>
<td>0</td>
</tr>
<tr>
<td>Emotivity</td>
<td>7</td>
<td>6.0</td>
<td>1</td>
<td>0.9</td>
<td>3</td>
</tr>
<tr>
<td>Total</td>
<td>44</td>
<td>37.9</td>
<td>9</td>
<td>7.8</td>
<td>9</td>
</tr>
</tbody>
</table>

Table 6.4 shows that pragmatic and deictic gestures (84.5% in total) dominate in the gesture use for Ellen, whereas the concrete iconic and metaphoric iconic gestures relatively rarely (15.6% as a whole) co-occurred with her progressive utterances. The prototypical gesture accompanying the progressive utterances denoting “continuous ongoingness” and “futurity” is the deictic gesture, whereas the typical gesture with those denoting “habituality”, “temporary state” and “emotivity” is the pragmatic gesture.

Compared with the results of GUESTS, you find that there are twice as many progressive utterances made by the GUESTS co-occurring with concrete iconic and metaphoric iconic gestures as there are made by ELLEN (34.2% vs. 15.6%). Similarly, only the progressive utterances with “continuous ongoingness” were more likely to be
accompanied by the concrete iconic gestures for both the GUESTS and ELLEN.

The findings suggest that only when the progressive construction has the “continuous ongoingness” construal can it be the most likely to co-occur with iconic gestures; otherwise, the co-speech gestures mainly have a pragmatic function related to the discourse. Thus, the progressive construction in the interview context did not effectively facilitate the native English speakers individually or as a community to focus on the event-internal structure. This is inconsistent with what previous linguists have claimed and as psychologists operating in labs have found.

6.3.2.2 Semantic elements for which the gestures with progressive utterances are iconic

As the sample of iconic (concrete and metaphoric) gestures is small (N = 60), I combined the totals from the GUESTS and ELLEN for analysis in this section. In addition, due to the rather scarce co-occurrence of iconic (concrete and metaphoric) gestures with the progressive construction denoting non-aspectual meanings by both the GUESTS and ELLEN, I left them out (N = 7) and took only the iconic and metaphoric gestures accompanying the progressive construction with temporal-aspectual use into consideration (N = 53). Consequently, there were 33 iconic (concrete and metaphoric) gestures accompanying progressive utterances with a “current ongoingness” construal, 6 accompanying those with a “habituality” construal, and 14 accompanying those with a “temporary state” construal. Fig. 6.1a presents the distributions of the semantic elements that they are iconic for. The gestures co-occurring with progressive construction with “continuous ongoingness” construal are the most iconic for the actions of events (39.4%), and then next most frequently for the path of events (24.2%). In addition, it is just the progressive construction with “continuous ongoingness” construal that is accompanied by gestures showing a combination of semantic elements (most are combinations of action and entity) involved in events (15.2%). Both the progressive utterances denoting “habituality” and “temporary state” are predominantly accompanied by gestures that are iconic for
entities participating in the events (50% for the former and 35.7% for the latter), and second most often by those that are iconic for the process of abstract events (33.3% for the former and 28.6% for the latter).

Fig. 6.1a Semantic elements of the iconic gestures with progressive construction

6.4 Discussion

In order to examine whether, and if so, how, the polysemy of a grammatical construction correlates with multimodality, I investigated the relationship between the co-speech gestures and the progressive construction with rich meanings in natural conversation.

There are three main findings from our study. First, the utterances with a progressive (be V-ing) construction had greater potential to be multimodal than to be monomodal, both for the general population and for the individual. This pattern appears to stem from the fact that the progressive utterances with “continuous ongoingness” were far more likely to be accompanied by gestures.

Second, the prototypical gesture type for the progressive construction in general was pragmatic gestures. For the general population, this is the case regardless of its meanings, whereas for Ellen as an individual with a special role in the conversations,
the prototypical gesture type with the progressive construction with the construal of “habituality”, “temporary state” and “emotivity” was pragmatic gestures, and with the construal of “continuous ongoingness”, “repetition” and “futurity” it was deictic gestures.

In terms of our third finding, two pieces of evidence appear to further clarify the view that it is the use of the English progressive aspect to express “continuous ongoingness”, rather than the progressive construction in general, that facilitates the speakers’ mental simulation of the internal structure of events. First, only when the progressive utterances denote “continuous ongoingness” did the speakers produce predominantly (concrete and metaphoric) iconic gestures. Second, only the iconic gestures accompanying the progressive utterances with a “continuous ongoingness” construal was more likely to be iconic for the action, the nucleus of an event, and then to be iconic for more than two event elements.

Therefore, I will discuss the results in detail in terms of the multimodality of the polysemous progressive construction, in general, and of the progressive aspect, in particular.

6.4.1 Polysemy of the progressive construction and multimodality

The dominant use of co-speech gestures accompanying utterances with the progressive construction provides further evidence for the general claim that grammar is (at least sometimes) multimodal in nature (Fricke 2008; Harrison 2009). Grammatical meanings are abstract and schematic; accordingly, there is no particular fixed gestural form for any particular meaning denoted by the progressive construction. Thus, the grammatical construction just correlates with gesture rather than with a particular kind of gesture.

The prototypical gestures for the utterances with the polysemous progressive construction were pragmatic gestures, and deictic gestures were also frequently involved. I believe such results derive from the inherent nature of the interview as involving a particular kind of spoken language usage event (Cienki 2015), that is, ones with a high level of interaction and involvement. Interviews, as face-to-face dialogue,
have particular social as well as semantic and syntactic features (Bavelas 1990; Bavelas et al. 1992). Pragmatic gestures refer directly to the interlocutor rather than to the topic, which serves to include the interactant (Bavelas et al. 1992). The native speakers of American English in the data make systemic use of pragmatic gestures to signal and maintain interaction and involvement not only with the interlocutor in conversations but also with the audience in the studio or the audience watching the show on TV.

The forms of the pragmatic gestures are relatively simple, predominantly with open hand gestures. As Fig. 6.1b shows, more than half of the pragmatic gestures were palm-up open hand (PUOH) gestures (58.2% for the guests and 64.8% for Ellen). Almost 15% of them were palm-down open hand gestures (PDOH). The guests used three times as many palm-lateral open hand gestures (PLOH) as Ellen did. The other forms, like extended vertical index finger, palm vertical open hand gestures, were 12.8% for the guests, and 14.8% for Ellen. However, the overall differences between Ellen and the guests were not significant ($\chi^2(3) = 2.465$, $p > 0.05$).

The PUOH gestures are also called “palm presenting” gestures (Kendon 2004: 273), which are assumed to be derived from the everyday action of presenting something, offering or giving something, or showing readiness to receive something via the extended open hand with palms up (Kendon 2004; Müller 2004). In our corpus, the native speakers of American English frequently made use of them to present to the
addressee or the audience what they were doing or are doing, the temporary state they were in, and attitudes or comments they had towards something as if on the empty palms of their open hands. In (2a), they were talking about the performance Gwen Stefani made in a video clip that had just been played on the screen. While explaining that it was not to be the video as Ellen said but only a live show, Stefani presented the fact of what was ongoing — “we were doing the live show” — on her open hands with palm up towards the addressee, Ellen (see Fig. 6.2a). In (2b), after seeing a picture in which the guest, Ricky Gervais, was running, Ellen used the progressive utterance with epistemic construal i.e., inferring the purpose of his running. As Fig. 6.2b shows, she extended her open hands with palms up towards the addressee, Ricky, while uttering “you are actually”, and the gesture held through “trying to get in shape”. Thus, she presents her inference via her PUOH gesture for Ricky’s agreement, which was signaled further by the next utterance “right?”). Unfortunately, Ricky denied her inference by saying “well, no.” and clarified that he worked out nearly every day to keep healthy instead of running only over a short period to keep in shape for participating in The Golden Globe Award show. Such PUOH gestures towards the addressee directly do not represent any information about the speech content but function to include the addressee to an extreme extent in interaction.

(2) a. Ellen: Because you were so raw in the video, too. You can just see how you are feeling throughout the whole—it’s just one take.

Gwen: I’m going to cry when I look at that. The thing is, that was never going be the video. That was just going to be -- [We were doing the show live], we were doing the song live for the first time in New York.

b. Ellen: [You are actually trying to get in shape] for the (.). It seems like you joke about it, but you are running, getting in shape for the globes, right?

Ricky: Well, no. Believe it or not, I work out nearly every day.
In contrast, the PUOH gestures towards the audience rather than the speaker in Fig. 6.2c and Fig. 6.2d tended to get the audience involved in their interaction. In (2c), Ellen and the guest, Charlie Puth, were talking about the process of getting in touch with Charlie and inviting him to appear on the Show. Charlie did not believe that he was able to be invited, so he took them as crank calls and hung up the phone every time. Consequently, the Ellen Show team had to look for him via his high school, etc. After Charlie’s narration of this, Ellen reidentified what they did with the progressive utterance “we were tracking you down”. Then Charlie confirmed her opinion and further reiterated it by saying “you trackers like were tracking my town down”. Meanwhile, he made a PUOH gesture towards the audience and turned his eye gaze from Ellen to the audience too (see Fig. 6.2c). Apparently, he presented the idea on his open hands with palms up to the audience to seek their agreement. Accordingly, the
audience replied with laughter in return.

(2) c. Charlie: ...I’m like, “Cool, Bye”. And then you called my high school, and you called my grammar school, like, everybody.

Ellen: We were tracking you down.

Charlie: You really did. [You trackers like were tracking my town down].

d. Ellen: I’m so inspired, I’m going to do that flagpole. I’m going to do it by the end of the season.

Kevin: [laughter] you are going to do flagpole?

Ellen: All right. [And now you are laughing at me. I’m going to do it.]

In (2d), Kevin was talking about getting stronger and stronger by doing flagpole exercises (i.e., a male fitness item for developing strong muscles) and tried to inspire more people to engage in this sport. Ellen responded that she was going to do it too, which stimulated Kevin Hart’s laughter and doubt about it. At the same time, the audience were clapping and laughing too. Therefore, Ellen made a progressive utterance “you are laughing at me” with the epistemic construal of re-identification of what Kevin and the audience were doing. Such a progressive utterance was accompanied by a PUOH gesture and also an eye gaze adjustment from Kevin to the audience in reply to the audience’s non-verbal involvement in the interaction too (See Fig. 6.2d).

The deictic gesture was the second most frequent gesture accompanying the
progressive utterance with various kinds of meanings except those denoting “current ongoingness”. Deictic gestures are generally taken to be employed by speakers to draw other people’s attention or to establish a joint focus of attention (Bavelas et al. 2014). In the corpus, the deictic gestures are usually in the form of extended index finger, the thumb and the palm up open hand pointed at the speaker, the addressee, the audience, the non-present person or object involved in the reported events, or those in the pictures or videos on the screen. As one of the most important characteristics of face-to-face dialogue, the pronouns, you and I in particular, were mainly the subjects of utterances. Thus, there are many progressive utterances that are accompanied by the deictic gestures of pointing at the speaker himself/herself or at the addressee to emphasize or clarify the agent of an event. The speaker also usually pointed at a certain object which was the shared common ground with the audience, as in (3a), when Ellen was questioning how the guest, Matt Lauer, could walk in the high-heeled shoes he had on. Matt explained it was not the first time he had done so and said to Ellen “you made a resolution and you came on my show”. Upon hearing this, the audience were laughed loud. Thus, Ellen reidentified the usage event that “They (the audience) are still laughing at your feet”, and she pointed at Matt’s shoe with her extended left index finger while uttering “they are still laughing at” (see Fig. 6.3a). The special female high-heeled shoes the male guest, Matt, was wearing caught the hostess’ and the audience’s attention the moment he stepped on the stage. The deictic gesture here just highlighted such shared common ground (cf. Holler & Wilkin 2009). In addition, the speaker can also make deictic gestures to point at abstract objects, such as in (3b), where Ellen asked what Leonardo DiCaprio did for Christmas. Leonardo told about the party he had with his family. In the previous gestures accompanying “they have loads of kids. Their kids had kids”, he set up the “family tree” model of family lineage in vertical tiers (successive generations below previous ones), using his open hands with vertical palm away from his body. While uttering the progressive utterance “their kids’ kids are having kids” with the construal of temporary state, he emphasized his point by pointing to the imaginary individuals and their kids on the previous “family tree” model he constructed in some downward sequences (see Fig. 6.3b). In this example, his pointing
to the kids may not only function to draw the listeners’ attention, but also to organize his own thoughts of how the generations flow.

(3) a. Matt: You made a resolution and you came on my show (..)

Ellen: I think they’re still laughing at your feet.

b. Leonardo: I have lots of brothers and sisters. And they have loads of kids. Their kids had kids, [and now their kids’ kids are having kids.]

Fig. 6.3a Gesture with “are laughing at your”

Fig. 6.3b Gesture with “are having”

Sometimes, the speakers profiled the setting of an ongoing event with deictic gestures. In (3c), Hilary Clinton, pointed at the backstage with her two separated index fingers to indicate to the audience the place where the event of her watching took place (see Fig. 6.3c). Furthermore, the deictic gestures can be used to clarify the information source which the speech did not provide, as in (3d), in which the girl, a member of the audience, explained why she had come to the show. In Fig. 6.3d, you can see that the girl sitting on the arm of the sofa pointed backwards at the screen with her right thumb when uttering “were saying like before” and held it through “you were single”. The boys said “they are single” in the video clips that played before the girl came to the stage. Therefore, the epistemic use of a progressive utterance was embodied in such a deictic gesture, citing the source of information.

(3) c. Ellen: You look rested and fantastic. I don’t know how that’s possible, but you do.

Hilary: You are so nice to say that. I was watching backstage]and I saw you talking
to Ashlyn and Brett.

d. Ellen: Why did you really come here?

    Audience: Well, I was (...) Okay, well, I’m a massive fan, Legit, but you guys were saying like before, you were single, and I’m single.

\[\text{Fig. 6.3c Gesture with “was watching backstage”} \quad \text{Fig. 6.3d Gesture with “were saying”}\]

### 6.4.2 Progressive aspect and multimodality

Previous linguistic and psychological comprehension studies stressed that the English progressive aspect should facilitate the focus on the internal structures of an event and stimulate the mental simulation of the ongoing process of an event. Although the prototypical gesture accompanying the utterances with progressive aspect is the pragmatic gesture, the progressive aspect-marked utterances in the data set with “continuous goingness” are most frequently accompanied by iconic (concrete and metaphoric) gestures. This result implies that for the American speakers in this real face-to-face interaction context, the English progressive construction was not being used with salient mental imagery that is specific to it. This could be because conceptual processing is highly flexible (e.g., Binder 2016; Dove 2016; Zwaan 2016), and the activation of grounded features varies across contexts. Relevant salient mental imagery was not strongly stimulated enough in the progressive aspect situations to give rise to
gestures or iconic gestures in the following contexts. First, when there were pictures or videos showing the past events in the immediate environment, the speaker tended to make no gesture or deictic gestures rather than iconic gestures, as in (4a) and (4b). As (4a) and Fig. 6.4a show, when the video on the screen is demonstrating the event that the little girl is acting out the animals, no manual gesture occurred. However, the eye gaze of the little girl, Brielle, her mother, and that of Ellen, converged on the video on the screen in front. In (4b), Ellen asked Charlie to explain his big kiss with Mechan at the end of his concert which was displayed in the video. When he said Mechan was his best friend and tried to explain something else, Ellen interrupted him, as she thought none of the audience was listening to him, because they were looking at the picture on the screen. Ellen pointed backwards at the picture on the screen with her right thumb when uttering “They are looking”, as in Fig. 6.4b. Meanwhile, the audience also started laughing.

(4) a. Ellen: You played “Head’s UP” with the animal noises.
        Brielle: Yes.
        Ellen: What animals did you (..), oh, you are acting it out.

b. Charlie: Anyway, so I made this record (..) Machan was (..)
        Ellen: No one is listening to you. [**They are looking**] at that.

![Fig. 6.4a](image1) No gesture with “are acting it out”  
![Fig. 6.4b](image2) Gesture with “are looking at”

On the other hand, it was only the progressive utterances denoting “continuous ongoingness” that were the most likely to be accompanied by (concrete and metaphoric)
iconic gestures. This gestural result provides further evidence that taking the progressive aspectual view reflects the speaker and enables the listener to view the event as ongoing dynamically (Comrie 1976; Madden & Zwaan 2003). The ongoingness and dynamicity of events apparently affords strong activation of the mental imagery of them. Consequently, the perceptual and motor simulations that underlie the speakers’ thinking and speaking are stimulated, which results in the representational gestures (Hostetter 2014; Hostetter & Alibali 2008). Most iconic gestures mimicked the action that the agent of an event executed or the movement that the speaker may have imagined in fictive events, as in (5a). When talking about the plastic dots around her mouth on a helmet that she wore during the film shooting, the speaker explained that the function of them was to capture her expressions. In fact, the dots as technical devices were attached to the helmet, recording her expression. However, the speaker alternately moved her two vertical hands with palms away from her body up and down twice while uttering “are capturing” and held it during “my expression” (see Fig. 6.5a). Thus, I can say her gestures reflected her mental simulation of the working mechanism of the plastic dots.

(5) a. Ellen: What do the dots do? What are they doing?

Nyong’o: [They are capturing my expression], so I had four cameras that were hooked to like a helmet on my head.

b. Ellen: But you are in crazy costumes and then behind the camera directing.

Eva: [laughter] I (..) we had a beach scene, so [I was running back and forth] in this outfit.
In addition, unlike the results of the previous studies on the effect of aspect on the gestural representations of motion events, the concrete iconic gestures with the progressive utterances were more complex than those with the non-progressive utterances, in that the former were more likely to encode the manner of the motion events, and the latter to encode the path of the motion events (Duncan 2002; Parrill et al. 2013). Our study showed that the iconic (concrete or metaphoric) gestures co-occurring with the progressive aspect marked utterances of motion events primarily only represented the path of concrete or fictive motion events, as in (5b) and (5c). In the concrete motion event of running back and forth in (5b), the guest, Eva Longoria, alternately moved her left extended index from right to left several times to represent the path “back and forth” of her running (see Fig. 6.5b). In the metaphoric events like (5c), the hostess, Ellen, commented on what the guest, Krasinski, did in a film by saying “you are running towards danger”. Meanwhile, she made a gesture that was also iconic to the fictive path of the metaphorical motion of “running towards”. As Fig. 6.5c demonstrates, Ellen moved her right hand with lateral palm towards her body quickly from right to left to depict the path. The reasons why (American) English speakers are more likely to profile the path rather than the manner of motion events in natural conversations may be as follows. First, as English is a satellite-framed language (Talmy 2000), path is the focus of motion events. According to the phenomenon of “manner modulation” (McNeill 2005: 201), the speakers of satellite-framed languages often omit manner in gesture and profile the direction of motion by gesturing about its path. Co-speech gestures with progressive aspect-marked utterances highlight the path to encode
the duration and ongoingness of a motion event. Second, there are many metaphoric events in natural conversations, the actions of which are less capable of being simulated directly (Bergen 2007, 2010). However, paths are salient for sequential mental scanning. Third, the acting out of a path demands less effort on the part of the hands than expression of the manner of the motion events. Thus, it fits more with the principle of the least effort in communication (Zipf 1949).

(5) c. Krasinski: I don’t know we would have the courage to do the same.
   Ellen: And you don't have to do that. [You are running towards danger]
   instead of away from it.

d. Bryan: So when we got to [the scene where we are shooting in the jail],
   we thought, this is going to be important to see a man completely, physically stripped of everything.

Fig. 6.5c Gesture with “are running towards”
Fig. 6.5d Gesture with “are shooting”

There were also a few utterances with progressive aspect accompanied by gestures which were iconic for entities involved in the events. When taking a close look at them, I found that most of them were iconic for the entities (concrete or metaphoric) of the setting in which the event occurred, as in (5d). As Fig. 6.5d shows, Bryan Cranston held his two hands with palms towards center in front to mold the scene or the jail, that is, the setting or location where the event of film shooting took place. The maintenance of the gesture may imply the long duration of the film shooting there. It implies that speakers “mentally simulated locations” and then profiled them in the gestures while
producing progressive aspect-marked utterances. This complements Liu & Bergen’s (2016) results that language comprehenders “did indeed simulate locations” (p. 181) in listening to progressive aspect-marked utterances, and this is also the case in language production. However, entities were the least profiled by the iconic gestures accompanying the progressive aspect-marked utterances denoting “continuous ongoingness” in conversations. This may have been caused by the rich use of pronouns (you or I) as the subject of utterances and the priority of profiling the action or the path left by motion in the progressive aspect-marked speech context.

6.5 Conclusions

This study investigated the multimodal representations of polysemous progressive construction utterances. The results showed that the gesture rate and the prototypical gesture did not significantly differ in the polysemous progressive utterances, but the rate and characteristics of (concrete and metaphoric) iconic gestures accompanying them showed differences. Utterances with the progressive aspect have greater potential to be multimodal, particularly when they denote “continuous ongoingness”. The prototypical gesture accompanying the utterances with progressive construction was pragmatic gesture, the second most frequent being deictic gesture, no matter what meaning it denoted. However, only when the progressive construction represented the “continuous ongoingness” of an event was the co-speech gesture more likely to be iconic, and predominantly iconic to the action or path of motion.

There are two clear implications of the findings. First, context plays the central role in determining the multimodality of the polysemous progressive construction in general. The context of talk show or conversation requires high interaction and involvement across the speaker, the addressees, and the immediate and potential audience. Those characterizations of context result in the prototypical pragmatic gestures accompanying the progressive utterances, regardless of the meanings they denoted. Reference to the speaker, the addressee, the audience, and the pictures or video
clips being talked about, in the immediate environment via deictic gestures, are also a critical reflection of the high interaction and involvement in the talk show. Without doubt, this needs further experimental study to compare the gestures produced in monologues like retelling stories and in dialogues, like retelling stories, and in dialogues, like semi-spontaneous conversations.

Second, it is the meaning rather than the form of a grammatical structure that determines the variation of its co-speech gestures. As a polysemous grammatical construction, the progressive utterances denoting different meanings significantly influenced the iconic gesture representations. Only when the progressive constructions denoted “continuous ongoingness” were they more likely to be accompanied by iconic (either concrete or metaphoric) gestures. According to the GSA framework (Hostetter 2014; Hostetter & Alibali 2008, 2010), representational gestures derive from the strong simulation of imagistic-motor information of an event. Thus, the “continuous ongoingness” denoted by the progressive construction doubtlessly facilitated the speakers’ mental simulation of the spatial or motor information of the events’ intermediate stage, the actions or the path of motion in particular. The other meanings denoted by the same grammatical structure may have made the internal stages of an event less accessible, or provided no spatial or motor information to be mentally simulated at all. Thus, the speakers’ gestural behaviors adapted to the meanings of a polysemous grammatical construction in an appropriate way. The strong sensitivity to grammatical meanings in speech and co-speech gestures supports the position not only that is grammar meaningful, but also that the meanings of a grammatical structure can relate to different kinds of cognitive processing by speakers, which can be seen in embodied simulations, in gesture.