

The multimodal marking of aspect: The case of five periphrastic auxiliary constructions in North American English

Jennifer Hinnell

hinnell@ualberta.ca

<http://orcid.org/0000-0002-4809-3512>

Department of Linguistics, University of Alberta

Published version: <https://doi.org/10.1515/cog-2017-0009>

Abstract

Cognitive linguistics (CL) has, in recent years, seen an increase in appeals to include multiple modalities in language analyses. While individual studies have incorporated gesture, gaze, facial expression, and prosody, among other modalities, CL has yet to completely embrace the systematic analysis of face-to-face interaction. Here, I present an investigation of five aspect-marking periphrastic constructions in North American English. Using naturalistic interactional data (n=250) from the Red Hen archive, this study establishes a multimodal profile for auxiliary constructions headed by one of five highly aspectualized verbs: CONTINUE, KEEP, START, STOP, and QUIT, as in *The jackpot continued to grow* and *He quit smoking*. Results show that gesture timing, the structure of the gesture stroke, and gesture movement type, are variables that iconically and differentially represent distinctive aspectual conceptualizations. This study enhances our understanding of aspectual representation in co-speech gesture and informs the ongoing debate within CL and construction grammar circles of what constitutes conventionalization, or what constitutes a construction (mono- or multimodal) (Bybee 2006, 2010; Feyaerts et al. 2017; Zima & Bergs 2017a).

Keywords: co-speech gesture, aspect, *Aktionsart*, multimodality, auxiliary constructions

Reference: Hinnell, J. (2018). Co-speech gestures reflect lexical and grammatical aspect. *Cognitive Linguistics*, 29(4). Advanced online publication. <https://doi.org/10.1515/cog-2017-0009>

1. Introduction

Meaning-making is a process that emerges in face-to-face interaction. It relies on a range of modalities – the speech stream, facial expressions, hand, head, and body movements, and is influenced by a variety of factors ranging from individual speaker preferences to syntactic, discourse, and pragmatic forces, and the interactional context, to name only a few. Co-speech gesture – the movement of the hands that is coordinated with the speech stream – has been shown to be related to abstract levels of language processing (Huette et al., 2014, Chu and Kita 2008). One such abstract domain is at the centre of the study presented here, namely the linguistic-conceptual domain of aspect, which captures how events unfold over time. This multimodal study examines aspect-marking verb constructions in North American English. It contributes to the current discussion within cognitive and usage-based linguistics, as well as in the field of gesture studies, that the study of linguistic meaning should be based on language as it occurs in everyday interactive contexts (Kemmer & Barlow 2000; Langacker 2000; Bybee 2006, 2010; Müller et al. 2013; Müller et al. 2014), which are, indeed, multimodal contexts (McNeill 1992; Kendon 2004; Zima 2017). Cognitive linguistics, and, more specifically, construction grammar, have a “conceptual openness to all levels of usage features” (Zima & Bergs 2017a: 1). As such, studies over the past few years have argued for the need to include in the focus of linguistic analysis kinesic (including prosodic) patterns (Schoonjans 2014; Debras 2017; Lanwer 2017; Zima 2017). The question that arises, then, is whether the construction – as the basic unit of language – should be considered to be multimodal.

Taking a corpus-based, quantitative approach, this paper presents a case study of five periphrastic auxiliary constructions in North American English and the manual gestures that are co-produced with them. Each construction comprises an aspect-marking auxiliary, i.e. a lexical verb

that has been co-opted to function as a ‘helping’ verb, in this case to mark aspectual force. The auxiliaries – CONTINUE, KEEP, START, STOP, and QUIT – are followed by a complement verb, as in the utterances (1) to (5) from the Red Hen multimodal archive used for this study:

- (1) Things start going good in life and they don’t think they deserve it.
- (2) My parents never said ‘Stop making up stories’.
- (3) I quit doing Botox about five years ago.
- (4) ... to find out whether or not it continues to grow.
- (5) The reason I kept getting married was I wanted to feel like I was normal.

I investigate to what extent each auxiliary construction exhibits a conventionalized gestural profile, and what this profile is. Principal research questions are: (1) Is the aspectual nature of the auxiliary verb expressed in co-speech gesture? (2) If so, in what way? That is, in which parameters of the gesture profile can aspectual construal of the event be observed? (3) Are form parameters hitherto not part of gesture form annotations, such as gesture co-occurrence, gesture/speech asynchrony, and segmentations of the gesture phrase, recruited for the marking of aspectual information? Importantly, the data stem from natural, unscripted interactions, and the 250 instances are analyzed quantitatively, an approach still relatively uncommon in gesture studies (Schoonjans 2014; Lanwer 2017; Zima 2017).

Only very recently have there been any corpus-based studies that have included a measure of the rate of co-occurrence of co-speech behaviour with particular utterances. Schoonjans (2014) reports 23%-53% correlation between German modal particles and certain forms of head movement. Zima (in press) compares spatial constructions with motion constructions and reports gesture co-occurrence of 37%-75% for English motion verbs, while Zima (2017) shows the gesture co-occurrence of the partially filled [all the way from X PREP Y] construction to be considerably higher; it is instantiated gesturally 80% of the time. Lanwer problematizes the issue of gesture co-occurrence and its relation to entrenchment as follows:

Defining what is meant by ‘in a sufficient number’ seems to be substantial if we want to determine the unit status of a given pattern in the light of frequency measurements. However, this only seems to be necessary if we want to establish a clear-cut distinction between patterns which are entrenched and patterns which are not.

(Lanwer 2017: 3).

Citing Langacker’s observation that “linguistic structures are more realistically conceived as falling along a continuous scale of entrenchment in cognitive organization” (Langacker 1987: 59), Lanwer concludes that “it seems to be unnecessary, if not impossible to define something like a frequency-threshold” (Lanwer 2017: 3). In their review of numerous studies, Feyaerts et al. similarly conclude:

The co-occurrence rates of verbal and gestural patterns range in frequency, depending on the perspective one takes, but they (obviously) never reach a 100 percent match. Where one needs to draw the line, however, still is very much a point of debate and is considered by some to be impossible to operationalize.

(Feyaerts et al. 2017: 147)

As Zima argues, however: “Empirical case studies on a wide range of constructions (and from different languages) [...] are the indispensable prerequisite for understanding the systematicity of constructions’ co-occurrence with manual gestures in multimodal language use” (Zima 2017: 1).

In light of these standpoints, and the paucity of studies with which to compare frequency data, I do not offer premature conclusions regarding the entrenchment of co-speech gestures with particular constructions. However, co-speech gesture is shown to be a frequent element of all auxiliary constructions examined here (overall, 59% of the instantiations were co-expressed with gesture). Further results show that the gestures corresponding to these auxiliary verb constructions conventionally mark the aspectual information of the auxiliary verb through gesture timing, the structure of the gesture stroke (i.e. number of segmentations),¹ and gesture movement type.

¹ Kendon (1980) considered a gesture phrase to be composed of four unique phases: preparation, stroke, hold, and retraction. The gesture stroke phase is “considered to be the most important gestural movement

In the discussion section, I return to the matter of frequency and co-occurrence, and evaluate the contribution of both the frequency results and findings for other variables vis-à-vis the call for support within linguistics, and cognitive linguistics specifically, that constructions be considered multimodal entities.

2. Aspect

The constructions investigated here are headed by aspect-marking auxiliaries. Aspect refers to how an event or process unfolds over time; it expresses whether an event is complete or ongoing, punctual or long in duration, and continuous or repeated, among other distinctions. Importantly, aspect allows a speaker to convey different temporal construals of the same event (Comrie 1976; Frawley 1992; Croft 2012). For example, both of the following two sentences depict the reading of a story in the past tense. However, in (6), the reading of the story is bounded and presumably complete, whereas in (7) the past progressive signals that the event is ongoing and imperfective.

(6) Charlie read a story.

(7) Charlie was reading a story.

English is a language known to have impoverished grammatical marking of aspect, i.e. it marks very few aspectual distinctions morphologically (Comrie 1976: 1). Its aspect-marking morphemes are limited to the ‘ing’ morpheme, which characterizes the progressive construction as in (7) above and denotes imperfective aspect, and the regular past tense marker ‘ed’, which, though it is a tense marker, also conveys perfective aspect. Instead, English recruits a wide range of grammatical and lexical constructions to convey aspectual information, from adverbial

phases as the stroke usually carries the gestural meaning, constitutes the nucleus of gestural units, and correlates with the most relevant part of the verbal utterance” (Bresse et al. 2013: 1102). It is normally the stroke that is annotated for hand form, movement type, direction, etc.

phrases (*again and again, over and over, at once*) and directional particles (*he ate it up, she was singing away, they drove on*; Rice and Newman 2004), to periphrastic verb constructions (e.g. *John started talking*), such as those featured in this study. The auxiliary verbs in these constructions verbs have moved along an ‘auxiliation path’ from their initial lexical meanings towards a grammaticized aspectual meaning that partially echoes the semantics of the original lexical root (Kuteva 2001; Heine & Kuteva 2002). That is, as auxiliaries, they influence the resulting construction with their *Aktionsart*² (also known as inherent aspect) to allow for a re-construal of the temporal structure of the original event. For example, the verb *sneeze* is inherently punctual and bounded, and refers to a single event, as in (8). The progressive in (9) can give a reading of Mac being in the middle of either a single sneeze or of a multiple-sneeze event. However, only the multiple event reading is available in (10).

- (8) Mac sneezed.
- (9) Mac was sneezing.
- (10) Mac kept sneezing.

Thus, in (10), the inherent aspect of the auxiliary *keep*, which means ‘to hold or retain’ in root form, imbues the *keep sneezing* construction with a stronger, iterative meaning (Talmy 2000b: 68; Heine & Kuteva 2002: 19).

Finally, aspect has been categorized into three major groupings: open, closed, and phase (Chung & Timberlake 1985; Brinton 1988; Frawley 1992).³ The distinction between open and phase aspect forms one of the organizing principles for this study. Open aspects include all those that *open* events, namely the imperfective, durative, progressive, and habitual. These aspects express the extension of an event over time such that the event is ongoing or happens routinely;

² *Aktionsart* refers to aspectual distinctions that belong intrinsically to the semantic meaning of a verb (Comrie 1976, Talmy 2000b).

³ Brinton (1988) uses imperfective and perfective to denote open and closed aspect respectively.

they focus on the internal time of the event and are unbounded (e.g. *Claire lives in Europe*). In the present study, CONTINUE and KEEP constitute open aspect auxiliaries. By contrast, *closed* aspect events are viewed as complete, unitized events, and are bounded.⁴ (There are no closed aspect constructions featured in this study.) Phase aspect completes the hierarchy of aspects and focuses on the end points of an event. It includes inceptive aspect (*She started to talk*), prospective aspect (*They are about to leave*), and terminative aspect (*She quit smoking*). In this study, START denotes inceptive aspect, and STOP and QUIT denote terminative aspect.

As discussed, open aspect focuses on the internal structure of an event. This raises the important point that there are different *ways* for an event to be open. An open aspect event can be structured as one continuous event, in which case an iterative reading means a repetition of the whole event. An open event can, alternatively, develop in phases, or incrementally. Talmy's term *plexity* (Talmy 2000b) is helpful in capturing this potential ambiguity in open aspect. Plexity captures a notion of quantity for both objects and for time. Should an object or event consist of only one element, it is considered *uniplex*, whereas an event or object that is complex in some way is referred to as *multiplex*. For example, in English the noun phrase *a bird* has a single referent, which can be 'multiplexed', or pluralized, by adding *-s* to the noun, as in *birds*. For a verb phrase such as the earlier example with *sneeze*, the verb phrase inherently predicates a single or 'uniplex' relation consisting of one sneeze event (i.e., a closed event that is bounded and highly punctual). However, this relation can be coerced into a multiplex interpretation by using it in the [KEEP VVG] construction, as in *She kept sneezing* (Talmy 2000b: 69). In the gesture annotations in this study, the variable I term *action phases* (described in the next section) correlates with

⁴ *Boundedness* can refer to a 'natural endpoint' in the case of an accomplishment or achievement, as in *write a letter*. In the case of grammatical aspect, and as used here, boundedness "refers to an action that is finished, whether it has a natural endpoint that has been reached, or simply terminates" (Croft 2012: 77).

plexity: gesture phrases that contain only one articulation in the stroke are considered uniplex gestures, while those containing two or more are considered to be multiplex.

While aspect captures the diverse ways in which events unfold over time, Talmy's (2000a) notion of force dynamics captures how entities interact with respect to the internal or external force that activates the event. Johnson (1987: 42) prompts us to consider that "our bodies are clusters of forces and that *every* event of which we are a part consists, minimally, of forces in interaction." Despite only beginning to receive attention in gesture research, force dynamics has been acknowledged to occupy "a central place in embodied approaches to meaning" (Mittelberg 2013: 772) and as such, needs to be considered in light of the target utterances of this study. Consider the minimal pairs given in (11) and (12). While the first sentence offers an objective construal of a ball rolling (the ball is unimpeded), in (12) the addition of the auxiliary *keep / keep on* invokes one of two possible force dynamic patterns: (1) the ball has a tendency to roll but is being impeded (e.g. long grass), or (2) the ball has a tendency to rest but is being propelled (e.g. by the wind). Thus the "kept on" connotes a countervailing force being overcome.

- (11) The ball was rolling along the green.
- (12) The ball kept (on) rolling along the green.

(Talmy 2000a: 412)

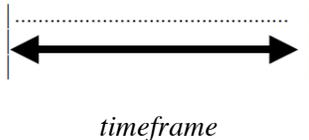
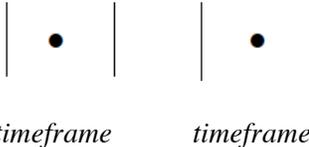
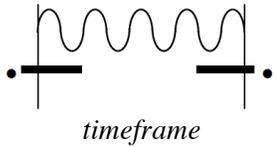
In this study, CONTINUE and KEEP differ in their force-dynamic profile in a similar way to these two examples, with CONTINUE being the more neutral of the pair, force-dynamically speaking.

Although STOP and QUIT are not characterized by Talmy, I suggest that they can also be considered a force-dynamic minimal pair: QUIT signals a more forceful termination of the event than STOP and seems to be associated with a higher degree of intention (under what Talmy refers to as the psychological realm). For example, in (13) and (14), the first utterance suggests that Charlie will come back to her homework later, while in the second there is an inherent finality to the event – the homework will remain unfinished.

- (13) Charlie stopped doing her homework.
 (14) Charlie quit doing her homework.

In Figure 1, I summarize the major types of aspect and force dynamics and their relation to the auxiliaries profiled in this study. The bold lines in the image schema diagrams mark the element that is profiled in each aspect. Thus, for open aspect, the straight line in the diagram excludes the endpoints of the event timeframe. For closed aspect, the dot in the schematic diagram signifies that it is the complete event that is profiled; a closed event has no internal structure. For phase aspect, the squiggly line denotes the timeframe of the event, while the bold markings denote the profiled onset or end of the event.

Figure 1. Characteristics of aspect, event structure, and target auxiliaries

Category of aspect	Image Schema	Characteristics	Auxiliaries
OPEN ASPECT imperfective habitual/durative progressive iterative/incremental		extension over time; modifiable, unbound- ed	CONTINUE, KEEP* <i>It <u>continued</u> to grow.</i> <i>She <u>kept</u> getting married.</i>
CLOSED ASPECT perfective punctual telic semelfactive		complete, unitized, bounded	(not profiled in this study) <i>Charlie <u>read</u>.</i> <i>Mac <u>ate</u> lunch.</i>
PHASE ASPECT inceptive terminative		change of status with- in or outside time frame of event; focus on endpoints: onset/ offset	START, STOP, QUIT* <i>Things <u>started</u> going well.</i> <i><u>Stop</u> making up stories.</i> <i>I <u>quit</u> smoking.</i>

* Denotes the force-dynamically strong auxiliary in each set.

2.1. Aspect embodied in gesture

Speech and gesture have been argued to comprise two components of a human's cognitive and communicative system (McNeill 1992; Kendon 2004). Concomitantly, mental representations

have been shown to be intimately bound to the workings of the human body (Gibbs Jr. 2005; Bergen 2012). Studies using the Action-sentence Compatibility Effect (ACE) (Glenberg & Kaschak 2002) have shown that language processing uses both motor and perceptual systems to perform mental simulations, and eye-tracking studies have suggested that more abstract levels of language processing are grounded in body movement (Huette et al. 2014). There is also strong evidence of a speaker's tendency to produce more gestures when talking about spatial information than about verbal or neutral topics (Alibali 2005).

In studies directly investigating aspect, it has been found that co-speech gestures “are part of a linguistic-conceptual representation in which the verb aspect has a role” (Duncan 2002: 183). In one of the few studies to include both grammatical and lexical aspect, Duncan investigated spoken and gestured expressions in Mandarin and English in an elicitation task to examine if differences in how aspect was encoded “might be expected to associate with differences in conceptual representation during acts of speaking” (2002: 190). Duncan found systematic, within-language co-variation of the choice of aspect and/or *Aktionsart* and the features of co-occurring gesture. For example, she found that when speakers produced imperfective-marked speech, their gestures were longer in duration – as if people were conceptualizing an ongoing event – and more complex (which Duncan describes as “temporally-extended, repeating or agitated movements” (2002: 183) compared to the gestures produced in perfective aspect speech contexts.

Grammatical differences in aspect have been shown to generate differences in manual action in the ACE experimental paradigm. For example, it's been shown that when speakers use sentences in the progressive aspect to talk about hand motions, this facilitates manual actions in the same direction; however, when they utter the same sentence but in the perfect aspect, the ef-

fect on manual actions is absent (Bergen and Wheeler 2010). Similarly, when people viewed videos of car accidents (Matlock et al. 2012), they gestured differently according to the aspectual force of the instructions they were given. Participants were asked what *had happened* (perfective framing) or what *was happening* (imperfective framing). Findings showed that “the form of aspect used in the question differentially influenced the way people conceptualized and described actions” (Matlock et al., 2012, p. 699). Aspect has also been shown to affect gestures accompanying event descriptions based on texts containing simple past or progressive marking (i.e. only grammatical aspect). Similar to Duncan, Parril et al. (2013) found that when participants described an event using progressive aspect marking, the description was accompanied by longer-lasting and more complex gestures (though only in the case when progressive aspect was also marked in the original text form).

Aspectual distinctions influence eye movement as well as gesture. Huette et al. (2014), compare the pattern of eye movements tracked while participants listen to the storytelling of events in the simple past and in the progressive. Eye-movement fixations in a simple past description (e.g. *He went*) showed the greatest density in the central region of a screen and saccades were of a longer duration (“as if staring at a static object or scene” (2014: p.2), while the past progressive showed a more diffuse distribution on the screen and shorter saccades, which the authors suggest reflect the dynamic nature of the event.

With the exception of Duncan (2002), most experimental studies of gesture and aspect restrict themselves to the distinction between the past progressive (*was driving*) and simple past (*drove*), typical English proxies for the nearly universal imperfective/perfective distinction. Furthermore, auxiliaries has been characterized as a highly imagistic process, in which auxiliaries “reflect general conceptualization capacities crucially involving imaginative, or rather *imaging*,

aspects of human cognition” (Kuteva 2001: 19). Yet, little is known about how auxiliary constructions with strong lexical aspect such as those featured here are gestured. This study aims to fill (some of) this gap in the literature.

3. Methods

3.1. Red Hen and target utterances

The study presented here used the Edge Search Engine of the Distributed Little Red Hen LabTM, an online multimedia archive whose core dataset consists of the NewsScape Library of International Television News (Steen & Turner 2013). At present, Red Hen consists of over 250,000 hours of audiovisual data from public broadcasts, representing circa 3 billion words of closed-captioned text, and grows daily. It contains a wide variety of genres, from news broadcasts to talk shows, comedy, and advertisements. For this study, only naturalistic and interactional contexts were included, meaning that scripted shows, prepared news reports, political speeches, and other non-interactional genres were excluded.

The auxiliaries targeted in this study occur in one of two constructions: the infinitival [AUX to VVI], as in *continue to go*, or the progressive [AUX VVG], as in *stop talking*.⁵ Whereas in English, KEEP, STOP, and QUIT are available only in the progressive construction, CONTINUE and START are available in both constructions (i.e., *start talking* and *start to talk*). I used frequency of construction type in The Corpus of Contemporary American English (COCA) to determine which construction to use for the study. A search of the spoken portion of COCA revealed that for CONTINUE, the infinitival [CONTINUE to VVI] was most frequent, while for START, the progressive [START VVG] was most frequent.

⁵ VVG and VVI are standard tags used in the British National Corpus (BNC) and COCA: VVG is the progressive inflected form and VVI is the infinitive verb form.

Finally, there are other verbs within a similar semantic frame that have undergone the same path of grammaticalization from main verb to catenative auxiliary, such as *begin*, *finish*, *carry on*. The target items included in this study were chosen over these near-synonymous counterparts due to their considerably higher overall frequency in the spoken register of the COCA corpus (e.g. spoken COCA featured 28,531 instances of [STOP VVG] and only 917 instances of [FINISH VVG]).

3.2. Data collection

The data were collected from TV programs broadcast between September 1, 2010, and September 1, 2013. The search string consisted of a two-part verb phrase: the possible inflected forms of the auxiliary (e.g. *keep/keeps/kept*) plus the ten most frequent collocating verbs for that auxiliary construction according to COCA. Search returns in Red Hen are given in chronological order starting with the most recent. The search returns were first viewed and processed to determine which constituted valid data points for this study. Valid clips featured interactional, unscripted discourse in which the speaker's hands were visible and unencumbered (e.g. not holding a microphone). The first 50 valid clips formed the dataset for each auxiliary, yielding a total of 250 clips.

This set of 250 videos was then viewed again for gesture co-occurrence. This measure of how frequently the expression is gestured speaks to the degree of conventionalization of gesture for each auxiliary. Gesture co-occurrence was coded as 'yes' if there was a new gesture at the time of the target utterance, and 'no' if there was no gesture, or if there was no change in gesture (i.e. the gesturer maintained a preceding gesture form as a rhythmic marker, or maintained a hold of a previous gesture form). Annotations were completed by the author, with 20% of the dataset

annotated by a second gesture coder. Interrater reliability for gesture co-occurrence was 90%. Of the 250 valid clips, 147 co-occurred with gesture.

3.3. Data annotation

The 147 instantiations of the target constructions that featured gesture, produced by 126 unique speakers, were downloaded from Red Hen for annotation in Elan. The approximate length per clip was 40 seconds – 15-20 seconds on either side of the target utterance. The variables and levels included in the gesture annotations are shown in Table 1. Gesture movement annotations are based on the Linguistic Annotation System for Gestures (LASG) (Bressem 2013; Bressem et al. 2013). Two variables – gesture timing and action phases – are not part of the LASG and are defined by the author.

In the LASG, gesture movement patterns are characterized by movement type, axis, and direction. Together the movement variables form the “movement trace –the lines and contours of objects and abstract concepts that leave “imaginary traces in gesture space” (Mittelberg 2008). When there was no appreciable displacement in space (e.g. a wrist turn), the data point was coded as ‘static’.

Table 1. Annotation schema

Variable	Levels
Movement type	straight, arced, circle, static ⁶
Movement axis	vertical, lateral, sagittal, static (e.g. wrist turn only)
Movement direction	up, down, left, right, towards body, away from body, static
Action phases	number of separate articulations within a gesture stroke
Hands	left, right, both
Gesture timing	timing of gesture onset prior to target auxiliary onset in msec

⁶ The LASG also features spiral as a movement type. Given that repeated arc forms can be difficult to distinguish from spiral gestures, both spiral and arc gestures were coded as arc in this study.

Action phase refers to the number of separate segments within the gesture stroke. For a cyclic gesture this aligns with the number of rotations.⁷ However, action phase can also capture segmentation within a straight stroke. Straight gestures have been described as having ‘singularity of form’ and being one unit (Cienki 1998). However, as the data in this study show, there is no reason to assume that gestures that proceed with a straight movement type are likely to be executed in one distinct phase. A segmented gesture stroke along the vertical axis is shown in Video Example (V1).

(V1)^{8,9}



G: continued to grow and grow

Speech tier: *Ever since then [fame] has just continued to grow and grow and grow.*
 Gesture tier: hands holding metaphorical object move in straight, vertical movement in five distinct action phases.

Iterative and incremental development over time are two ways in which open (unbounded, progressive/durative) events can unfold. *Action phase* captures repeated rotations of the arm within the same cyclic gesture stroke, indicating repetition or iterativity; but it can also capture a phased expression of a straight gesture trajectory, as shown in V1 and associated with incrementality.

⁷ Ladewig (2011) annotates for the number of rotations in a cyclic gesture, however, these and other studies seem to capture this only for curved movement forms, i.e. spiral, circle, and arc.

⁸ Screenshots from Red Hen will be referenced with V (video) to distinguish them from speech examples. Red Hen files are cited in the file format: date (yyyy-mm-dd), broadcast time, network, program, and timestamp in seconds.

⁹ 2012-07-16_2300_US_KNBC_The_Ellen_DeGeneres_Show_669-719

Iterative and incremental development over time are two ways in which open (unbounded, progressive/durative) events can unfold. The number of *action phases* is closely tied to the conceptual category of *plexity* described above. A singular circle gesture or a single-phase straight gesture would constitute uniplex events, while any gesture with an action phase greater than one was considered to be multiplex.

Lastly, to capture gesture timing, the target utterance was considered to begin directly at the onset of the auxiliary, regardless of upstream predicates in the verb phrase. For example, in (15), the onset was measured from “continue”, rather than from the beginning of the inflected verb phrase:

(15) We’re going to have to continue to be smart about the way we do business.

The beginning of the preparation phase in the gesture was used to calculate onset timing. The LASG relies on a “frame-by-frame marking procedure” based on Seyfedinnipur (2006) to establish the onset time of a gesture and the same procedure was followed here. The timing of the onset of both the gesture and the target utterance was recorded in milliseconds. The difference between the gesture and speech onsets gave the asynchrony, with a negative value indicating that gesture preceded the related utterance.

Interrater reliability was 92% for gesture form variables. Disagreements resulted primarily from varying use of the ‘static’ annotation and were resolved once the definition was clarified. Any remaining disagreements were settled by consensus. In Table 2, I summarize the predicted associations between the variables and the features of aspect introduced in this section.

Table 2. Predicted relationships between gesture variables and aspect type

Variable	Predictions
Movement type	open aspect: more cyclic gestures to reflect ongoing process. phase aspect: more straight gestures to indicate event onset/offset.
Axis/direction	open aspect: horizontal to reflect event unfolding over time. phase aspect: vertical to reflect focus event onset/offset.
Action phases	open aspect: 2+ APs to reflect multiplex event (iterative or incremental); 1 AP to reflect uniform extension. phase aspect: 1 AP to reflect uniplex onset/offset of event. No multiplex strokes: event extension is not available in phase.
Gesture timing	open aspect: longer onset to reflect ongoing event. phase aspect: shorter onset to reflect focus on event boundary.

4. Findings

To give an impression of the gestures that characterize each auxiliary, I first give a qualitative account of each construction, including syntactic and semantic characteristics, and a prototypical gesture profile. In section 4.2 I provide quantitative results for each auxiliary and a statistical comparison by open and phase aspect.

4.1. Semantic overview by auxiliary

4.1.1. Open aspect auxiliaries

CONTINUE

The dataset for the [CONTINUE to VVI] construction was characterized by sentences such as:

- (16) If a star fish ends up getting an arm cut off, it continues to grow.
- (17) I think we will continue to see a downward slope.
- (18) The big story continues to be North Korea.
- (19) This is one group of people that continues to have to change.

The most frequent complement verbs were *be* (with a token frequency of 16), *have* (7), *do* (6) and *grow* (4). Third person subjects were dominant (82%) and the vast majority (90%) of the ut-

CONTINUE also featured several instances of cyclic movements. These consisted of multiple stroke segments, as exemplified in (V3). Here, the movement is a smooth circle trace that is iterated twice to indicate an uninterrupted event progression.

(V3)¹¹



G: whether or not it continues to grow

Speech tier: ...*to find out - ah - whether or not it continues to grow*.

Gesture tier: two cyclic movements with right hand and return to rest position on knee.

KEEP

The KEEP construction was characterized by sentences such as (20) to (23). Again, as exemplified in these utterances, the majority were abstract in nature. Frequent complement verbs were *say* (14), *go* (12), and *get* (10), and there was near equal distribution of first person (40%) and third person (38%) inflection.

(20) So when there is no winner, that jackpot keeps getting higher.

(21) But the potential candidates just keep coming.

(22) I keep saying it's heartbreaking.

(23) It's like god said, maybe this once I'm just going to keep going.

KEEP prototypically consisted of repeated straight gestures on the vertical axis and repeated cyclic gestures. In (V4) and (V5) the same speaker presents each of these forms over the span of the same utterance. In the first, the speaker produces a cyclic movement iterated four times for “keep coming back”. Here the rotation is counter-clockwise or leftward, which is unusual for a right-handed cyclic gesture. This appears related to the semantics of “coming back”; gestures

¹¹ 2013-06-03_1200_US_FOX-News_Fox_and_Friends_1658-1708.

indicating past time have been shown to be gestured leftward on the speaker's horizontal axis (Casasanto & Jasmin 2012). The speaker then produces a multiplex gesture as shown in (V5). The hand shape is a point, and this form is repeated four times downwards on the vertical axis aligned with "keeps getting paid".

(V4)¹²



G: keep coming back

Speech tier: : ...*because if you keep coming back you keep getting paid.*

Gesture tier: cyclic movement counterclockwise; iterated four times.

(V5)¹³



G: keep getting paid

Speech tier: ...*because if you keep coming back you keep getting paid.*

Gesture tier: iterated downward stroke with pointed hand shape; four action phases.

¹² 2013-08-25_0700_US_FOX-News_Huckabee_2168-2208.

¹³ Ibid.

While the vertical axis overall prefers downwards gestures, (V6) shows a gesture for KEEP that moves upward on the vertical axis. The movement type was coded as spiral (arc) given the gesture's displacement in space. Both the direction and the incrementation – there are six action phases – iconically reference the metaphorical growth of the jackpot to a “higher and higher” number.

(V6)¹⁴



G: keeps getting higher and higher

Speech tier: *So when there is no winner, that jackpot keeps getting higher and higher.*

Gesture tier: upward moving bimanual arc (spiral) gesture; six action phases.

These examples show the variation within the gesture profile of the open aspect auxiliaries CONTINUE and KEEP. I now outline the semantic and gestural profiles for phase aspect auxiliaries.

4.1.2. Phase aspect auxiliaries

START

The inceptive phase aspect of START imbues the [START VVG] construction with a conceptual focus on event onset. The utterances for START (indeed, for all phase aspect constructions) tend to be more concrete than those of CONTINUE and KEEP, as illustrated in the following sentences from the dataset. The most common complement verbs were *talk* (12), *get* (9), *go* (5) and *do* (5), and the construction preferred inflection in the third person (44%).

¹⁴ 2013-08-07_1200_US_FOX-News_Fox_and_Friends_1589-1629.

- (24) Everything just started creeping up on me.
 (25) I started doing all these cores (core exercises).
 (26) And then he started talking about my family.
 (27) The water started coming over the edge.

Uniplex cyclic gestures, as well as uniplex straight gestures on the vertical and sagittal axes, characterize the gesture profile for START. An example of a cyclic gesture for START is shown in (V7).

(V7)¹⁵



G: start

making

blessings

Speech tier: *When the pope comes on Twitter [...] and says 'I'm going to start making blessings on Twitter'....*

Gesture tier: cyclic movement outwards on sagittal axis; one action phase.

STOP and QUIT

Sample utterances for the two remaining phase aspect auxiliaries, STOP and QUIT, both of which express terminative aspect, are given in (28) to (35), below. The most common complement verbs for STOP were *do* (11), *talk* (9), *use* (6), and *work* (5), while for QUIT they were *smoke* (20), *drink* (8), *do* (6), and *use* (5). [QUIT VVG] is the only auxiliary construction to prefer the first person (42%): *quitting* appears to be something that one speaks of doing oneself. STOP, by com-

¹⁵ 2013-08-07_1500_US_KNBC_Today_Show_2098-2158

parison, is more often used in the third person, or as command (second person imperative). Like START, STOP and QUIT both featured more concrete than abstract events.

- (28) Nobody really wants to stop doing this.
- (29) If you own one, stop using it and contact them for a free repair.
- (30) Maybe you should just stop talking altogether ...
- (31) 500,000 people stopped looking for jobs.

- (32) I quit doing Botox about five years ago.
- (33) And you began with the first major decision, to quit drinking at the age of 40.
- (34) And so he's told the republican candidates to quit using the word capitalism.
- (35) Our loudest opponents on the left are never gonna like us so let's quit trying to curry favour with 'em.

QUIT was the only auxiliary to prefer lateral strokes. The profile that emerges is of a two-handed, lateral gesture, exemplified below in (V8): a sweeping sideward gesture made with both hands moving outwards in a flat hand shape. This form has been reported previously in the literature as the 2-palm-down-across and is characteristic of utterances which contain negative force (Harrison 2013). It features a relatively higher velocity (qualitatively observed) in its stroke and an abrupt end to the stroke phase. Importantly, in instances when this bilateral stroke occurs with multiple action phases, the segments are incremental, i.e. the stroke is segmented into phases along one outward lateral trajectory, rather than the gesture stroke returning to its starting point for a second outward motion. The gestures produced with the [STOP VVG] construction also display this outward moving, bimanual gesture pattern, as shown in (V9), though the end of the stroke phase was qualitatively observed to be less abrupt in its execution.

(V8)¹⁶G: quitsmoking

Speech tier: *I lit a cigarette [...] and put it out, like, I'm gonna quit smoking.*

Gesture tier: bimanual outward gesture on lateral axis.

(V9)¹⁷G: stoppedtalking

Speech tier: *Remember when I stopped talking to you for 6 months?*

Gesture tier: bimanual outward gesture on lateral axis.

In this section, I have shown a range of contexts of use of the five auxiliary constructions and the gesture forms that accompany them. It is clear that some form features are related to the semantic representation of individual events, in addition to the aspectual construal of the event. In the next section, I abstract away from these individual differences and present quantitative and statistical findings.

¹⁶ 2013-01-03_0735_US_KNBC_Tonight_Show_with_Jay_Leno_3235-2350

¹⁷ 2013-06-04_1600_US_KABC_Live_With_Regis_and_Kelly_1576-1626

4.2. Quantitative findings

Quantitative findings are presented by aspect type, beginning with open aspect auxiliaries followed by phase aspect auxiliaries. Table 3 gives an overview of the results for each auxiliary, with open aspect verbs in columns on the left and phase aspect verbs on the right.

Table 3: Summary of results for each auxiliary

Variable	Open Aspect		Phase Aspect		
	CONTINUE	KEEP	START	STOP	QUIT
Gesture co-occurrence					
absolute freq. (n=50)	37	29	25	27	29
relative freq.	74%	58%	50%	54%	58%
Gesture timing (mean)	-386 msec	-356 msec	-269 msec	-293 msec	-171 msec
Action phases (mean)	2.1	2.6	1.5	1.5	1.6
Movement axis ^{&}	vertical (47%)	vertical (61%)	vertical (46%)/ sagittal (46%)	vertical (56%)	horizontal (52%)
Movement type ^{&}	straight (75%)	straight (64%)	straight (63%)	straight (88%)	straight (74%)
Hands (preferred hand and relative frequency)	1 hand (62%)	1 hand (52%)	1 hand (60%)	2 hands (52%)	1 hand (62%)

[&] most frequent, relative frequency given in brackets

4.2.1. Open aspect

CONTINUE

With a gesture co-occurrence rate of 74%, the [CONTINUE to VVI] construction is the most gestured auxiliary. Although a chi-squared test yielded no significant differences between gesture co-occurrence across auxiliaries, a Pearson residuals posthoc test does show the observed rate of gesture for CONTINUE to exceed the expected rate. CONTINUE also features the greatest mean asynchrony between the onset of the gesture phrase and the onset of the target utterance (-386 msec). Table 4 shows the distribution of movement type by axis, plexity, number of action phases and gesture timing for CONTINUE. The distribution of axis by movement type is signifi-

cant ($p < .05$).¹⁸ A post hoc test of Pearson residuals reveals that both straight gestures on the vertical axis and arced/circular gestures on the sagittal axis occurred more frequently than expected (these cells are noted in bold in the table).

Table 4. [CONTINUE to VVI] – Summary table (n=37)

	Axis			Plexity		APs	Asynchrony
	Lateral	Sagittal	Vertical	Uniplex	Multiplex		
Arced	2	2	0	2	2	1.8	-435 msec
Circle	0	4	1	0	5	3.8	-1.06 sec
Straight	5	6	16	13	14	1.9	-259 msec
Static	n/a	n/a	n/a	1	0	1	-230 msec
TOTAL	7	12	17	16	21	MEAN 2.1	-386 msec

* bold indicates significant cells

While circle gestures are predominantly multiplex – indicating iteration, straight gestures are almost equally distributed between uniplex and multiplex strokes. These findings indicate a fairly even distribution of aspectual structure: straight uniplex gestures show a smooth (i.e. unsegmented) extension over time such as in (V2) above, while the multiplexed straight gestures, as in (V1), provide additional aspectual detail, in that case an incremental unfolding of the event.

KEEP

With a gesture co-occurrence rate of 58%, KEEP is less gestured than CONTINUE. The distribution of axis, plexity, mean action phase and onset timing by movement type are shown in Table 5.

KEEP features an interaction of axis and movement type for (chi-squared test, $p < .005$); a Pearson residuals posthoc test confirmed that straight movements on the vertical axis and arced movements on the lateral axis occur more frequently than expected.

¹⁸ To avoid 0-cells in the chi-squared test, arced and circle were reduced to a non-straight category here.

Table 5. [KEEP VVG] – Summary table (n=29)

	Axis			Plexity		APs	Asynchrony
	Lateral	Sagittal	Vertical	Uniplex	Multiplex		
Arched	2	1	1	1	3	3.3	-217 msec
Circle	2	3	1	1	5	2.3	-679 msec
Straight	0	3	16	7	12	2.5	-283 msec
TOTAL	4	7	18	9	20	MEAN 2.6	-356 msec

* bold indicates significant cells

As Table 5 shows, all gestures for KEEP display a strong preference for multiplex. This suggests that gestures for KEEP are more frequently used to add additional information about the internal aspectual structure – i.e. whether an event develops iteratively or incrementally – than simply depicting the unboundedness of the event in a uniplex stroke.

4.2.2. Phase aspect

The three remaining catenative auxiliaries – START, STOP, and QUIT – are phase aspect auxiliaries, that is, they mark movement into or out of the time frame of the event.

START

With a gesture co-occurrence rate of just 50%, [START VVG] has the lowest incidence of gesture across the five auxiliary constructions. Table 6 shows the distribution of movement type by all variables. There is no significant interaction of movement type by axis, however, given the general preference across the other auxiliaries for vertical strokes, the relatively higher frequency of movement along the sagittal axis for START is worth noting. START also features a higher frequency of cyclic gestures (24%) than either of the open aspect auxiliaries.

Table 6. [START VVG] – Summary table (n=25)

	Axis			Plexity		APs	Asynchrony
	Lateral	Sagittal	Vertical	Uniplex	Multiplex		
Arced	1	2	0	2	1	1.3	-300 msec
Circle	0	4	2	2	4	2.2	-377 msec
Straight	1	6	9	13	4	1.3	-224 msec
TOTAL	2	12	11	16	9	MEAN 1.5	-269 msec

STOP

The terminative, phase aspect construction [STOP VVG] featured 27 instances (54%) that were aligned with gesture. Results are summarized in Table 7. There is an interaction of axis by movement type ($p < 0.05$); the Pearson residuals posthoc test shows that the cell-wise difference was in straight movements on the vertical axis, which, again, were dominant. Of the 8 lateral movements, 5 feature both hands moving outward (as shown in (V9) above). Of note is that STOP is the only auxiliary not to feature any cyclic gestures.

Table 7. [STOP VVG] – Summary table (n=27)¹⁹

	Axis			Plexity		APs	Asynchrony
	Lateral	Sagittal	Vertical	Uniplex	Multiplex		
Arced	2	1	0	1	2	1.7	-73 msec
Circle	0	0	0	0	0	n/a	n/a
Straight	6	2	14	13	9	1.5	-332 msec
Static	n/a	n/a	n/a	2	0	1.0	-200 msec
TOTAL	8	3	14	16	11	MEAN 1.5	-293 msec

* bold indicates significant cells

QUIT

QUIT features a gesture occurrence rate of 58%, making it the most gestured of the three phase aspect auxiliaries. As shown in Table 8, movement on the lateral axis is dominant, representing a marked difference from other auxiliaries, and indicative of the profile shown in example (V8).

¹⁹ Two instances had no movement, thus the axis columns total 25 rather than 27.

QUIT also features a higher proportion of uniplex strokes than all other auxiliaries, at 72%. Furthermore, removing the outlier (a circular motion with six APs, which I return to in the discussion section below) would result in QUIT having the lowest mean action phases per stroke, at 1.4 APs.

Table 8. [QUIT VVG] – Summary table (n=29)²⁰

	Axis			Plexity		APs	Asynchrony	
	Lateral	Sagittal	Vertical	Uniplex	Multiplex			
Arced	6	0	0	4	2	1.3	-150 msec	
Circle	0	1	0	0	1	6.0	-300 msec	
Straight	8	4	8	15	5	1.5	-332 msec	
Static	n/a	n/a	n/a	2	0	1	-35 msec	
TOTAL	14	5	8	21	8	MEAN	1.6	-171 msec

4.3. A statistical comparison across auxiliaries

I now offer statistical results across open and phase aspect auxiliaries. That is, taken together, I examine how the results for open aspect (CONTINUE and KEEP) compare with the results for phase aspect (START, STOP and QUIT). I also discuss other interactions in the data that require attention.

There were two significant effects by aspect type, as shown in Table 9. Interactions were determined by one-way ANOVA. Though there is no effect of gesture co-occurrence by aspect, there is a main effect of action phase by aspect ($F(1,144)=18.86, p<.0001$) and a main effect of gesture asynchrony by aspect ($F(1,145)=4.21, p<.05$). These findings indicate that open aspect gestures are reliably produced further in advance of their target utterances, and feature more action phases, than their phase aspect counterparts.

²⁰ Two instances had no movement, thus the axis columns total 27 rather than 29.

Table 9: Interactions by aspect type

	Open aspect	Phase aspect
Action phases [*]	2.4 APs	1.5 APs
Timing [§]	373 msec	243 msec
Gesture co-occurrence [^]	45%	55%

^{*}p<.0001, [§]p<.05, [^]not significant

There were no significant effects of axis and direction by aspect. However, when compared across auxiliaries rather than aspect type, there were significant correlations. A chi-squared test showed an effect of auxiliary on axis ($p<.001$) and a Pearson residuals posthoc determined that QUIT has a preference for the lateral axis and an aversion to the vertical axis, while START prefers the sagittal axis and has an aversion to the lateral axis. Thus, although gestures across all auxiliaries most frequently features movement along the vertical axis as shown in the summary tables above, axis does appear to play a unique role in the profiles for START and QUIT.

An analysis of gesture direction contributes further information to the auxiliary profiles. A chi-squared test of gesture direction by auxiliary was significant ($p<.05$). A Pearson residuals posthoc showed that QUIT and STOP both featured right-left movements (bimanual gestures with each hand moving outwards), and KEEP had a greater than expected occurrence of upwards movements on the vertical axis.

The distribution of open and phase aspect by gesture movement type (straight, circle, or arc) is shown in Table 10. While there was no significant effect, relative frequency shows that straight gestures dominate for both open and phase aspect, while circle gesture traces are more characteristic of open aspect. Arc gestures are more frequent for phase aspect auxiliaries.

Table 10: Relative frequency distribution of movement type by aspect

	Open aspect	Phase aspect
Arced	12%	15%
Circle	17%	9%
Straight	68%	70%
Static	3%	6%

Interactions not related to aspect type include an effect of movement type on both gesture timing and mean action phases, as shown in Table 11. This table shows that circular movements have the most asynchronous onset timing and the greatest number of action phases. However, it is worth noting that the mean onset asynchrony for circle gestures with open aspect auxiliaries CONTINUE and KEEP is -870 msec, while for phase aspect (START and QUIT; STOP had no circular gestures) it is -338 msec, which supports the finding of aspect type on gesture timing. The predictions stated at the outset were that the phase aspect auxiliaries would uniformly feature a more synchronous onset given their aspectual emphasis on the beginning or end of an event, which is indeed the case, even when the interaction of movement type on onset timing is taken into consideration.

Table 11: Movement Type by mean action phases and gesture timing

	Mean APs [*]	Gesture timing [§]
Arced	1.9	-230 msec
Circle	2.9	-660 msec
Straight	1.8	-260 msec

*p<.05, §p<.001

In sum, the findings show that type of aspect plays a role in both plexity (number of action phases) and gesture timing. These are the two variables that were predicted to most abstractly represent the aspectual contour of an event. Since the auxiliaries CONTINUE and KEEP have an inherent aspect (*Aktionsart*) that ‘multiplexes’ an event (Talmy 2000b), it was predicted that the open aspect gesture constructions would involve gestures with more action phases within

the stroke when compared to the inherently ‘uniplexing’ semantics of the phase aspect constructions START, STOP and QUIT. Results show this prediction to be borne out. Phase aspect auxiliary constructions prefer one AP, while the open aspect constructions are characterized by two or more APs. The movement variables – type, axis and direction – did not correlate significantly with aspect type, but did yield significant results within auxiliaries. These findings speak directly to the existence of a prototypical gesture form for each auxiliary.

5. Discussion

In this study, I have moved beyond the parameters traditionally considered to be the loci of meaningful content in gesture and included stroke segmentation into action phases and gesture onset timing in an examination of aspectual constructions. Results show that both gesture timing and action phases are variables that reliably, and I suggest, iconically, represent aspectual information in a multimodal manner. The qualitative and quantitative data reveal that each aspectual auxiliary construction has elements that are unique to it, and that the type of aspect – open or phase – differentiates the gestural profiles as well. Here, I discuss possible semantic roots of the profiles that emerged, and conclude with a discussion of how this study provides support for considering these multimodal constructions.

Firstly, given the central place of force dynamics in embodied meaning (Mittelberg 2013), it befits us to consider these findings in relation to the force-dynamic pairs identified earlier. As discussed, CONTINUE and KEEP are a semantic minimal pair with regard to their force-dynamic characteristics: CONTINUE signals an unencumbered continuation of an event, while KEEP conveys that an exertion of force is required to shift an object or event into motion or for it to continue in motion. Here, CONTINUE has a longer onset timing at -386 msec compared to -356 msec for KEEP. I propose that the correlation of longer onset with CONTINUE could be due to the

semantics that signal an unencumbered continuation vis-à-vis KEEP. For the action phase variable, the higher mean number of action phases for KEEP (2.6) vs. CONTINUE (2.0) may iconically represent the additional effort required to maintain the motion of the event, which the gesturer performs (though not consciously) through continued stroke iteration.

The force-dynamic profile of STOP and QUIT form a similar minimal pair; [STOP VVG] refers to a neutral cessation of an event while [QUIT VVG] is inherently more forceful. The findings showed that, while STOP can be both uniplex (59%) and multiplex (~ 30% had 2 APs), QUIT significantly prefers a uniplex gesture. QUIT also has the shortest onset of all auxiliaries, at -171 msec before the auxiliary, compared to -293 msec for STOP. The more synchronous onset and single action-phase gestures for QUIT vs. STOP could be accounted for in the force-dynamic contrast between these auxiliaries. Though a direct link is difficult to ascertain in a corpus study, the impact of force-dynamics on the findings presented here certainly requires further investigation.

The study of gesture is inherently complex in part due to the fact that “gestures convey meaning in different ways than speech does”, namely, in a global and synthetic process rather than a linear process (Zima & Bergs 2017a: 4). Thus, although core forms do emerge, there are outliers, such as the cyclic gesture with six action phases for QUIT shown in (V10). This example represents a notable departure from the dominance of bimanual, outward gestures prototypical of QUIT.

(V10)²¹G: for me to quit talkingSpeech tier: *You were waiting for me to quit talking.*

Gesture tier: female speaker executes left-handed cyclic gesture with six action phases.

In this instance it is not the *quitting* event that is gestured, but rather the *persistence of the talking*. This affects both movement type and number of action phases and points to the ability of the speaker/gesturer to underline – through co-speech gesture – a specific element of a construal.

In her seminal work on cyclic gestures, Ladewig (2011) showed that depictive cyclic gestures (which she differentiates from word-search gestures) refer to an ongoing process “in every instance” (2011: 6). Thus, for this study, cyclic gestures were expected to be most common in open aspect events, CONTINUE and KEEP. However, the finding that was not expected was the high proportion of cyclic forms with START constructions. Ladewig proposes that there is a metaphorical extension of the image schema CYCLE that invokes the motion of a crank as a process in a machine, as in starting up an engine. The over-representation of cyclic gestures with START could be related to the gesture metaphorically representing the start, or ‘cranking up’ of a process, rather than the ongoing nature of the process itself. This, again, speaks to the ability of the speaker/gesturer to choose which aspect of an event to highlight in his/her construal.

²¹ 2013-02-08_0735_US_KCBS_Late_Show_with_Dave_Letterman_1840-1890

The profiles presented here represent an aggregate of the gesture forms. For each auxiliary there are instantiations that are close to the core profile, and instantiations that vary widely from it. Some gestures occurring with open aspect constructions have one action phase, while occasionally a gesture with a phase aspect construction features three APs. Indeed, the major operational problem with multimodal data lies in the multitude of variations that are possible. As outlined in the literature cited earlier on frequency and co-occurrence, and further elucidated in Mittelberg (2007), a person's gesture is influenced by a multitude of factors, including individual idiosyncrasy. In addition to noted individual preferences in gesture space, there are many contextual factors that may influence gesture that have only begun to be examined and which, furthermore, are difficult to control for in naturalistic data. These include linguistic variables, such as inflectional biases (Rice & Newman 2005, 2008; Hinnell & Rice 2014), and concrete vs. metaphorical usage reported here. Further complicating matters is the way that gesture can be used to highlight different salient elements of a construal in a specific context of use, as in the cyclic gesture for QUIT shown above.

This brings us to the major issues of frequency and recurrence that are central to the discussion of what constitutes a multimodal construction. This study provides an empirically grounded analysis of patterns that emerge for each auxiliary related to aspectual profiles, and provides evidence to support a preliminary constructional profile for these five auxiliary constructions. With regard to the matter of frequency and co-occurrence as markers of constructionhood, no single study can resolve this issue. Indeed, in light of Langacker's view of entrenchment as continuously scalar (Langacker 1987), it may not be possible (Lanwer 2017), or necessary (Zima 2017), to determine a precise threshold frequency as a measure of its entrench-

ment.²² Certainly the field as a whole is wrestling with this issue and further empirical studies are needed. However, the study presented here contributes evidence of a high degree of conventionalization of the gestural profiles associated with these five aspectualized auxiliary constructions.

6. Conclusion

This study has shown that the five periphrastic auxiliary constructions examined here are differentially expressed in co-speech gestures and that the aspectual construal is iconically depicted in a range of gesture parameters. Results showed a degree of co-occurrence of 50-74%. Open aspect auxiliaries (CONTINUE and KEEP) reliably correlate with longer onset timing and a greater mean number of action phases per stroke, while phase aspect (START, STOP and QUIT) are correlated with more synchronous onset of gesture and fewer stroke segmentations. By including variables such as gesture co-occurrence, timing, and number of action phases, in addition to form variables, a more complete gesture profile emerges for each of these constructions. Correlations were also seen for certain auxiliaries and particular movement directions and movement types. These were shown to parallel semantic distinctions in both aspectual and force-dynamic characteristics of the event construal.

Cognitive linguistic and usage based approaches put speech at the centre of linguistic analysis. Increasingly, this means considering face-to-face events as the locus of meaning-making, which renders the object of study inherently multimodal. Monomodal research has, in recent years, proposed an expansion of the notion of construction to include phenomena that traditionally have not been included in the linguistic description of an utterance, such as viewpoint and discourse phenomena (Dancygier & Sweetser 2005; Dancygier 2008; Langacker 2008). The

²² For a full discussion of the different “criteria for constructionhood”, see Zima and Bergs (2017) and papers therein.

findings presented here support other recent studies suggesting that constructions should be considered multimodal units (Steen & Turner 2013; Schoonjans 2014; Debras 2017; Zima & Bergs 2017b). While many more studies are required to investigate degrees of entrenchment, the very fact that conventionalization between gesture and linguistic form can be shown suggests that these co-speech gestures ought to be considered a part of inherently multimodal constructions.

Although co-speech gestures are not an obligatory part of the aspect-marking auxiliary constructions investigated here, when they are gestured, they certainly exhibit conventionalized forms. While the ultimate question remains as to what degree of conventionalization should satisfy a cognitive linguist, here I have shown that aspectual gesture constructions do belong to the set of conventionalized gestures that are coordinated with linguistic processes. As Zima and Bergs (2017b) note, far more empirical studies, with considerably more depth and breadth with regard to linguistic features, gesture form, and contextual variables, are required. Given the emphasis of seminal studies to date on the multimodal representation of stance and pragmatic markers (e.g. Debras 2017, Schoonjans 2014), further studies that explore abstract and grammatical elements of construal, such as aspect, are needed.

Acknowledgements

I would like to thank Sally Rice for valuable feedback throughout this project. Earlier versions of this paper were presented at the International Society for Gesture Studies (San Diego, 2014), International Cognitive Linguistics Conference (Edmonton, 2013), and Association Française de Linguistique Cognitive (Lille, 2013), and I am grateful to those audiences for their responses. I would also like to thank three anonymous Cognitive Linguistics reviewers for their insightful comments and suggestions. This work was supported by a J. A. Bombardier Canada Graduate Scholarship (#767-2012-2405) from the Social Sciences and Humanities Research Council of Canada (SSHRC). I also gratefully acknowledge the Killam Trust for their assistance.

References

- Alibali, M. 2005. Gesture in spatial cognition: Expressing, communicating, and thinking about spatial information. *Spatial Cognition & Computation* 5(4). 307-31.
- Bergen, B. 2012. *Louder Than Words*. New York: Basic Books.
- Bergen, B. & Wheeler, K. 2010. Grammatical aspect and mental simulation. *Brain and Language* 112(3). 150-58.
- Bressem, J. 2013. A linguistic perspective on the notation of form features in gestures. In *Body – Language – Communication: An International Handbook on Multimodality in Human Interaction*, ed. by C. Müller, A. Cienki, E. Fricke, S.H. Ladewig, D. McNeill & S. Teßendorf, 1079-98. Berlin: De Gruyter Mouton.
- Bressem, J., Ladewig, S.H. & Müller, C. 2013. Linguistic annotation system for gestures. In *Body – Language – Communication: An International Handbook on Multimodality in Human Interaction*, ed. by C. Müller, A. Cienki, E. Fricke, S.H. Ladewig, D. McNeill & S. Teßendorf, 1098-124. Berlin: De Gruyter Mouton.
- Brinton, L. 1988. *The Development of English Aspectual Systems*. Cambridge: Cambridge University Press.
- Bybee, J. 2006. From usage to grammar: The mind's response to repetition. *Language* 82(4). 711-33.
- . 2010. *Language, Usage and Cognition*. New York: Cambridge University Press.
- Casasanto, D. & Jasmin, K. 2012. The Hands of Time: Temporal gestures in English. *Cognitive Linguistics* 23(4). 643-74.
- Chu, M. & Kita, S. 2008. Spontaneous gestures during mental rotation tasks: Insights into the microdevelopment of the motor strategy. *Journal of Experimental Psychology* 137(4). 706-23.
- Chung, S. & Timberlake, A. 1985. Tense, Aspect, and Mood. In *Language Typology and Syntactic Description: Grammatical Categories and the Lexicon*, ed. by T. Shopen, 202-58. Cambridge: Cambridge University Press.
- Cienki, A. 1998. STRAIGHT: An image schema and its metaphorical extensions. *Cognitive Linguistics* 9(2). 107-49.
- Comrie, B. 1976. *Aspect: An Introduction to the Study of Verbal Aspect and Related Problems*. Cambridge: Cambridge University Press.
- Croft, W. 2012. *Verbs: Aspect and Causal Structure*. Oxford: Oxford University Press.
- Dancygier, B. 2008. The genitive in constructional blends. In *New Directions in Cognitive Linguistics*, ed. by V. Evans & S. Pourcel, 161-81. Amsterdam: John Benjamins.
- Dancygier, B. & Sweetser, E. 2005. *Mental Spaces in Grammar: Conditional Constructions*. Cambridge: Cambridge University Press.

- Debras, C. 2017. The shrug: Forms and meanings of a compound enactment. *Gesture* 16(1). 1-34.
- Duncan, S. 2002. Gesture, verb aspect, and the nature of iconic imagery in natural discourse. *Gesture* 2(2). 183-206.
- Feyaerts, K., Brône, G. & Oben, B. 2017. Multimodality in Interaction. In *The Cambridge Handbook of Cognitive Linguistics*, ed. by B. Dancygier, 135-56. Cambridge: Cambridge University Press.
- Frawley, W. 1992. *Linguistic Semantics*. New York: Routledge.
- Gibbs Jr., R.W. 2005. *Embodiment and Cognitive Science*. New York: Cambridge University Press.
- Glenberg, A.M. & Kaschak, M.P. 2002. Grounding language in action. *Psychonomic Bulletin & Review* 9(3). 558-65.
- Harrison, S. 2013. The temporal coordination of negation gestures in relation to speech. Paper presented at the *TiGeR*, Tilburg, NL.
- Heine, B. & Kuteva, T. 2002. *World Lexicon of Grammaticalization*. Cambridge: Cambridge University Press.
- Hinnell, J. & Rice, S. 2014. Inflection-specific gestural constructions in English: The case of the catenative auxiliary. Paper presented at the *International Society of Gesture Studies*, San Diego, USA.
- Huette, S., Winter, B., Matlock, T., Ardell, D. & Spivey, M. 2014. Eye movements during listening reveal spontaneous grammatical processing. *Frontiers in Psychology* 5. 1-7.
- Johnson, M. 1987. *The Body in the Mind. The Bodily Basis of Meaning, Imagination, and Reason*. Chicago, IL: University of Chicago.
- Kemmer, S. & Barlow, M. (eds) 2000. *Usage-Based Models of Language*. Stanford: CSLI Publications.
- Kendon, A. 2004. *Gesture: Visible Action as Utterance*. Cambridge: Cambridge University Press.
- Kendon, A. 1980. Gesture and speech: Two aspects of the process of utterance. In *Nonverbal Communication and Language*, ed. by M. Ritchie Key, 207-88. The Hague: Mouton.
- Kuteva, T. 2001. *Auxiliation: An Enquiry into the Nature of Grammaticalization*. Oxford: Oxford University Press.
- Ladewig, S.H. 2011. Putting the cyclic gesture on a cognitive basis. *CogniTextes* 6. 1-23.
- Langacker, R.W. 1987. *Foundations of Cognitive Grammar: Theoretical Prerequisites*. Stanford: Stanford University Press.
- . 2000. A dynamic usage-based model. In *Usage-Based Models of Language*, ed. by S. Kemmer & M. Barlow, 1-63. Stanford: CSLI Publications.
- . 2008a. *Cognitive Grammar: A Basic Introduction*. Oxford: Oxford University Press.
- . 2008b. Metaphoric gesture and cognitive linguistics. *Metaphor and gesture*, pp.249-251.

- Lanwer, J. 2017. Apposition: A multimodal construction? The multimodality of linguistic constructions in the light of usage-based theory. *Linguistics Vanguard* 3(1). 1-12.
- Matlock, T., Sparks, D., Matthews, J.L., Hunter, J. & Huette, S. 2012. Smashing new results on aspectual framing. *Studies in Language* 36(3). 700-21.
- McNeill, D. 1992. *Hand and Mind: What Gestures Reveal about Thought*. Chicago: University of Chicago Press.
- Mittelberg, I. 2007. Methodology for multimodality. One way of working with speech and gesture data. In *Methods in Cognitive Linguistics*, ed. by M. Gonzalez-Marquez, I. Mittelberg, S. Coulson & M.J. Spivey, 225-48. Amsterdam: John Benjamins.
- . 2008. Peircean semiotics meets conceptual metaphor: Iconic modes in gestural representations of grammar. In *Metaphor and Gesture*, ed. by A. Cienki & C. Müller, 115-54. Amsterdam: John Benjamins.
- . 2013. The embodied mind: Cognitive-semiotic principles as motivating forces in gesture. In *In Müller et al.*, 750-99.
- Müller, C., Cienki, A., Fricke, E., Ladewig, S.H., McNeill, D. & Bressemer, J. (eds) 2014. *Body – Language – Communication: An International Handbook on Multimodality in Human Interaction* (Handbooks of Linguistics and Communication Science 2). Berlin: De Gruyter Mouton.
- Müller, C., Cienki, A., Fricke, E., Ladewig, S.H., McNeill, D. & Teßendorf, S. (eds) 2013. *Body – Language – Communication: An International Handbook on Multimodality in Human Interaction* (Handbooks of Linguistics and Communication Science 1). Berlin: De Gruyter Mouton.
- Parrill, F., Bergen, B. & Lichtenstein, P. 2013. Grammatical aspect in multimodal language production: Using gesture to reveal event representations. *Cognitive Linguistics* 24(1). 135-58.
- Rice, S. & Newman, J. 2004. Aspect in the making: a corpus analysis of English aspect-marking prepositions. In *Language, Culture and Mind*, ed. by M. Achard & S. Kemmer, 313-27. Stanford: CSLI Publications.
- Rice, S. & Newman, J. 2005. Inflectional islands. Paper presented at the *9th International Cognitive Linguistics Conference*, Yonsei University, Korea.
- . 2008. Beyond the lemma: Inflection-specific constructions in English. Paper presented at the *American Association for Corpus Linguistics*, Provo, Utah.
- Schoonjans, S. 2014. *Modalpartikeln als multimodale Konstruktionen: Eine korpusbasierte Kookkurrenzanalyse von Modalpartikeln und Gestik im Deutschen*: KU Leuven.
- Seyfedinnipur, M. 2006. *Disfluency: Interrupting Speech and Gesture*. Nijmegen, NL: Radboud University Nijmegen.
- Steen, F. & Turner, M. 2013. Multimodal construction grammar. In *Language and the Creative Mind*, ed. by M. Borkent, B. Dancygier & J. Hinnell, 255-74. Stanford, CA: Centre for the Study of Language and Information.
- Talmy, L. 2000a. *Toward a Cognitive Semantics, Vol. I*. Cambridge: MIT Press.

—. 2000b. *Toward a Cognitive Semantics, Vol. II*. Cambridge: MIT Press.

Zima, E. 2017a. On the multimodality of [all the way from X PREP Y]. *Linguistics Vanguard* 3(s1). 1-12.

—. 2017b. Multimodal constructional resemblance. The case of English circular motion constructions. In *Constructing families of constructions*, ed. by F. Ruiz De Mendoza, A. Luzondo & P. Pérez-Sobrino. Amsterdam: John Benjamins. 301-337.

Zima, E. & Bergs, A. 2017a. Introduction: Multimodality and construction grammar. *Linguistics Vanguard* 3(1).

—. 2017b. Towards a multimodal construction grammar. [Special issue] *Linguistics Vanguard* 3(s1). Berlin: De Gruyter Mouton.