English multimodal motion constructions. A construction grammar perspective.

Elisabeth Zima

Alexander von Humboldt-Foundation[1] / Freiburg Institute for Advanced Studies (FRIAS)

1. Introduction[2]

In a short commentary paper on current cognitive linguistic research on the relation between "metaphor and gesture", Jürgen Streeck (2008: 259-264) tells a short anecdote from a family conversation at the dinner table. One of the sons is trying to get the turn for a telling. Twice within a few seconds, he opens his turn by "to be honest" which is accompanied by a seemingly unremarkable gesture, i.e. the gesture is simply an uncurling of the fingers that were previously closed to make a loose fist, used by the owner of the hand as a headrest, holding his chin. The speaker opens his hand briefly, palm to the side, so that the fingers vaguely point in the direction of the audience. Streeck reasons that because the interactional sequence is just a single case, it is difficult "to make any claims about the relationship between form and sense. We can not know whether this gesture is the speaker's idiosyncracy or available as a 'prefab', 'rented' (Bakhtin, 1986) from the community" (ibid: 260f.). Given his micro-ethnographic and conversation analytical background, the issue of whether the gesture, in combination with the verbal pattern, is an entrenched and conventional routine, is probably not the central concern of Streeck's and other CA-oriented multimodality research. But it should matter to cognitive linguists and even more so, to construction grammarians.

---

[2] This paper gives just a rough overview of some of the findings of a more elaborated project on multimodal constructions. More detailed accounts are currently under review and in press, respectively (Zima, submitted and in press).

Construction Grammar (CxG) is concerned with the linguistic symbols that constitute our language system. Despite the fact that many cognitive linguists seem to be increasingly willing to support a view of language as also involving gesture, posture, facial expressions and other forms of bodily behavior, the focus of Construction Grammar has almost exclusively been on purely verbal structures. Since its earliest formulations in the late 1970ies/beginning 1980ies and most notably since Goldberg's (1995) groundbreaking dissertation on argument structure, an ever growing body of work has argued for specific verbal patterns to have idiosyncratic formal and/or semantic-pragmatic properties, which suggests that they are stored as distinct form-meaning pairs (or constructions) of speakers' linguistic system. Of the pioneers of the Construction Grammar framework[3], to the best of my knowledge, only Ron Langacker explicitly puts into question this status of constructions as (purely) verbal symbols:

> Manual gestures […], facial expressions, actions performed more globally (e.g. a shrug), and even factors like body language [...] may all be closely bound up with linguistic expressions, in which case they can hardly be excluded from 'language' on an a priori basis. […] When a baseball umpire yells *Safe!* and simultaneously gives the standard gestural sign to this effect (raising both arms together to shoulder level and then sweeping the hands outward, palms down), why should only the former be analyzed as part of the linguistic symbol? Why should a pointing gesture not be considered an optional component of a demonstrative's linguistic form?
> (Langacker 2008: 250).

However, Langacker is cautious to stress that not every gesture that we encounter in speech (cf. Streeck's example) automatically has linguistic unit status. Rather, "a structure per se qualifies as an element of a language just to the extent that it is entrenched in the minds of speakers and conventional in the speech community" (Langacker 2008: 250). In other words, for co-speech gestures to be reasonably assumed as part of a construction (both its form and its meaning), the combination of a verbal pattern and a given gesture has to be entrenched as a unit in the minds of speakers and conventional in the speech community.

Construction Grammar considers at least two main factors to play crucial roles in the individual entrenchment and socio-cultural conventionalization process: 1) recurrence, and 2) idiosyncracy. The former refers to the assumption that when speakers are confronted with a particular linguistic structure over and over again,

---

[3] Recently, quite a few studies have started to look at constructions from a multimodal point of view but just a few of them explicitly frame their findings within Construction Grammar (Andrén 2010, Steen & Turner 2013, Schoonjans, in prep).

they start to perceive it as a unit, a relatively fixed package of form and meaning which they store in their memory. Idiosyncracy, on the other hand, refers to specific formal and/or semantic-pragmatic properties that are inherent to a specific construction, i.e. they cannot be attributed in a compositional manner to its lexical components[4]. Taking this to the multimodal level, as suggested by Langacker, one has to assume that the same criteria, i.e. recurrence and/or idiosyncracy, apply. The following case study from the semantic domain of motion events focuses on the first criterion and is hence meant to provide empirical evidence for the claim that co-speech gestures can indeed be a recurrent feature across instantiations of constructions, more specifically of English motion constructions.

2.   Motion events

Over the past 15 years, the semantic domain of motion events has received a considerable amount of attention from linguists and psycholinguists. Their common root is Talmy's seminal work on motion events (1985, 2000) and his typological distinction of verb-framed and satellite-framed languages. This distinction is based on a typological survey on how *Path of motion* is lexicalized in the languages of the world. Take for instance English, a language of the satellite-framed type:

(1)   He ran out. (Nicoladis & Brisard 2000)
(2)   The rock slid/bounced/rolled down the hill. (Talmy 2000: 27f.)

In both examples, the *Manner of motion* is encoded in the main verb (running, sliding, bouncing, and rolling). The *Path of motion* is encoded in a satellite, i.e. an adjunct to the verb such as "out" or "down the hill"[5]. Verb-framed languages behave the other way round. They encode Path as part of the verb and Manner in an (optional) satellite. This holds for e.g. Romance languages such as French and Spanish.

(3)   Il est sorti en courant (He got out running, Nicoladis & Brisard 2000)

---

[4] The probably most famous example of such an idiosyncratic construction is the CAUSED MOTION-construction, studied in great detail by Goldberg (1995). Accordingly, verbs like "sneeze" or "shout" do not normally encode motion. However, used within examples of the caused motion-pattern such as "He sneezed the napkin off the table" or "He shouted him out of the room", they acquire a transitive motion-reading.
[5] This is not to say that there are no exceptions to that general pattern. Not the least loanwords from Romance languages like "exit", and "arrive" do encode Path of motion.

(4) La botella entró   a la cueva   (flotando).
     The bottle MOVED-in the cave (floating).
     (Talmy 2000: 49ff.)

More important to the current study than the typological distinction as such is Talmy's semantic close reading of motion events that guides his typology. He proposes that motion event descriptions may involve four main aspects[6]:

  (1) Figure: the moving object
  (2) Ground: the reference object that the figure moves relative to
  (3) Path: the path followed by the Figure
  (4) Manner: the specific way in which the Figures moves/is moved

The following case study focuses on Path and Manner of motion. These two aspects and their relation to co-speech gesture have recently received a great deal of attention, mainly in psycholinguistic research. The predominant methodology is to have people watch a cartoon where the protagonists are engaged in some motion action. Participants are then asked to retell the story to the experimenter or some other third party.

McNeill & Duncan (2000) found that speakers of English, where Manner is usually encoded as part of the verb, use gestures to depict Manner only when they put focus on the Manner-aspect, i.e. when it is a salient aspect of the analytical unit or "growth point" (McNeill 1992). Kita & Özyürek (2003) focus on how language structure influences the structure of accompanying gestures. Their main finding concerns the way speakers describe a scene in which Sylvester, from the Looney Tunes, swings from one side of a street to the other by hanging on to a rope. English speakers mainly used the verb "to swing" to describe the movement and used an arc-like gesture with it. Turkish and Japanese lack a verb such as "to swing," and accordingly speakers of these languages used verbs that did not encode arc-like movement. Regarding gesture, however, the picture is more blurred with speakers using flat, horizontal gestures and arc-like gestures alike. Hickmann, Hendriks & Gullbergs' (2011) focus is on the difference between English and French multimodal motion events. They show that English adults conflate Manner and Path in speech more often than French adults do. While both groups gesture mainly in reference to Path, English adults also conflate Manner and Path into single gestures.

These findings are relevant for the following case study as they shed light on motivations for adult English speakers to use iconic gestures as parts of motion event descriptions. Rather than conducting experiments to elicit multimodal data

---

[6] This list of four main aspects has been significantly enlarged and refined since its first proposal. FrameNet, for example, lists no less than 7 core- and 14 non-core components.

of motion event descriptions, the approach that is presented in the remainder of this paper is, however, based on naturally occurring language data from a wide range of discursive genres. For that matter, I conducted corpus research on the currently biggest video corpus for English: the Red Hen Lab Corpus. This web-based video corpus is currently being developed at UCLA under the direction of Francis Steen and Mark Turner[7]. It contains approximately 200 000 hours of video data from television broadcasts of all sorts (news broadcasts and shows, talk shows, late night shows, commercials etc.). The videos are linked to transcriptions so that the corpus is searchable for lexical strings[8]. Furthermore and in contrast to the aforementioned studies, the focus of this paper is not on all motion descriptions encountered in a given storytelling event, but on four specific motion constructions: [V(motion) *in circles*], [N *spin around*], [*zigzag*], and [*all the way from* X PREP Y]). The aim was to zoom in on the role of co-speech gesture[9] as accompanying these constructions. The following section presents a short summary of the case study's results.

3.   A case study on multimodal motion constructions

3.1. [V(motion) in circles]

This case study is based on 202 instances of the semi-lexicalized construction [V(motion) *in circles*]. Examples were extracted from Red Hen and assembled in a separate database if and only if they fulfilled two criteria: 1) the verb used within the construction is a motion verb (like go, swim, fly; not e.g. talk)[10]; 2) arms and hands of the speaker are visible on screen. All 202 examples were coded for the following parameters: (1) iconic Manner/Path gesture[11]: yes or no,

---

[7] I am very grateful to Francis Steen and Mark Turner for granting me access to their database.

[8] For more information go to https://sites.google.com/site/distributedlittleredhen/.

[9] Note that this case study focuses on hand gestures only. This restriction is necessary because of space limitations and to reduce complexity at this early stage of investigation.

[10] However, also metaphorical instantiations of motion verbs were included, as in (1) "at first he tried to go around in circles and still lie about it, and then I basically said at this point I just want to know answers" (Red Hen, file: CNN Starting Point, January 25, 2013 at 4:00 am).

[11] The satellite „in circles" encodes Path of motion. The verb refers to the Manner of motion. All gestures found to be clearly semantically related to the motion event description were iconic gestures that depict circular motion (Path). However, some gestures may also conflate Manner and Path aspects as e.g. a rapid circular hand movement accompanying "running in circles". Gestures that seemed to be related to the rhythm of speech (beats), as well as deictic gestures and hand movements that seemed to have a cognitive function, i.e. to help with word retrieval, were not coded as a case of a recurrent co-speech gesture that may potentially be an entrenched part of the construction. Their occurrence seems to be motivated by in situ metadiscursive or cognitive demands rather than by a close potentially entrenched semantic relationship with the construction under scrutiny. The absolute number of these instances of metadiscursive or cognitive gestures, however, is very small. Most

(2) semantic meaning of the construct: literal motion, metaphorical use, ambiguous use[12], (3) shape and orientation of palm and fingers, (4) use of one or both hands, (5) gestural depiction of one or more (incomplete) circles, (6) type of motion verb, and (7) discourse genre. For reasons of space, the following analysis concentrates only on aspects (1), (2), (3), and (7).

Table (1) shows that 60.4% of the 202 instances of [V(motion) *in circles*] were accompanied by an iconic path gesture, i.e. a circular movement of the hand(s).

	V(motion) *in circles*	
	gesture	no gesture
absolute	122	80
relative	60.4%	39.6%

Table 1: Number of instances of [V(motion) *in circles*], with or without gesture

However, the use of co-speech Manner gesture is not uniformly spread across all the construction's semantic uses. Table 2 shows that whether speakers do or do not co-depict circular motion by gestural means strongly depends on whether the construction is used in a literal or metaphorical sense. Descriptions of literal, i.e. physical, motion events incite speakers to gesturally depict the Path of motion in almost 69% of cases. In contrast, 63% of the metaphorical uses are not accompanied by circular gestures[13].

	literal		metaphorical		ambiguous	
	gesture	no gesture	gesture	no gesture	gesture	no gesture
absolute	99	45	13	22	10	13
group internal	**68.75%**	31.25%	37.14%	**62.86%**	43.48%	56.52%

Table 2: Number of instances according to the semantic use of the construction[14]

---

instances that were coded as "not containing gesture" did not contain hand movements of any kind. Also, I did not find examples, in which speakers encoded Manner but not (circular) Path.

[12] Coding of semantic uses is of course a non trivial thing to do. In search for a consistent method to differentiate between literal and metaphorical uses, I made use of the coding methodology developed by the Pragelazz group (2007).

[13] At first sight, this finding may seem little surprising, but recent psycholinguistic research on the processing of metaphoric language (Wilson & Gibbs 2007, Richardson & Matlock 2007, Johannson Falck & Gibbs 2011) has shown that speakers do activate and mentally simulate physical action even for metaphorically intended motion utterances. It therefore seems an interesting avenue for future research to explore how mental simulation relates to the active use of motion gestures in interaction.

[14] A chi-squared test (not taking into account the ambiguous cases) confirms that the use of the construction significantly influences the use of co-speech gesture, with literal uses favouring the use of co-speech gesture (Chi-square: 12,011, df: 1, p= 0.0001).

Which forms do these observed gestures have and do they qualify as "a recurrent commonality" (Langacker 2001:146)? The most obvious commonality is that the hand(s) or sometimes the hand(s) together with the forearm are performing a circular motion (on cyclical gestures, see also Kendon 2004 and Ladewig 2011). The hand shape is variable[15] but generally speaking, quite consistent. By far the most frequent iconic gesture, with 82 instantiations (67.21%), involves one hand, with the index finger being extended, the palm closed, and the wrist and index finger performing a circular movement (cf. the drawings from Red Hen examples, Fig. 1 and 2)[16]. In only eight cases, the extended index finger gesture is carried out using both hands. In 32 cases (26.23%), the gesture is an open, extended palm gesture, with the hand either held horizontally (palm up or down) (cf. Fig. 3) or vertically (Fig. 4). Again, the overwhelming majority of this type of circular gesture is carried out by one hand. Only in one single case, both hands are used.

Fig. 1: extended index, pointing down   Fig. 2: extended index, pointing up

Fig. 3: flat palm down, all fingers extended   Fig. 4: palm vertical, fingers loosely bent

As mentioned earlier, Red Hen is a corpus of TV broadcasts and contains a wide variety of discourse genres. Table 3 links discourse genre to the occurrence of circular path gestures. It suggests that with this particular construction the discourse genre has only minor influence on the use of iconic path gestures. Whereas the more spontaneous formats like talk shows and late night talks incite speakers to gesture rather freely and therefore to instantiate the [V(motion) *in*

---

[15] Two gestures are never exactly the same. They may vary to different degrees in hand shape, duration, velocity, placement in gesture space and many more parameters. Nevertheless, they can resemble each other in some aspects or have a core feature (like circular motion) which is shared across many instances.

[16] In 32 cases, the extended index finger is pointing downwards. In 33 cases, it is facing upwards, and in 17 cases, it is held horizontally.

circles]-construction together with an iconic path gesture slightly more often than expected (a quotient higher than 1 indicates more Path gestures than expected, below 1 less than expected), the more formal genre of news reports slightly discourages speakers to do so. This suggests that the gestural usage of [V(motion) *in circles*] is not a peculiar usage feature of some particular TV genre. Rather, it seems to be a conventionalized usage feature, instantiated by very different speakers in very different situations and conversational circumstances. This genre-independent pervasiveness may point towards its entrenchment as a multimodal form-meaning pairing (cf. conclusions).

genre		relative % in selection	circular gestures from that genre	quotient
talk show/late night talk		47.52%	54.1%	1.14
interview	sitting[17]	11.88%	13.11%	1.10
	standing	9.41%	8.2%	0.87
news report/news show	sitting	14.85%	9.83%	0.66
	standing	6.44%	4.92%	0.76
weather report		3.47%	5.74%	1.66
other (speech, trial,…)		6.44%	4.1%	0.64

Table 3: Occurrence of iconic circular gestures according to different discourse genres

The following sections present similar analyses for three other constructions: [N *spin around*], [*zigzag*], and [*all the way from* X PREP Y].

3.2. [N spin around]

Table 4 gives the overall distribution for the use of iconic Path/Manner gestures, i.e. hand movements that depict the circular movement expressed by "spinning around". Accordingly, in 72% of cases, speakers co-express the Path (and Manner) aspect of the construction gesturally. As with [V(motion) *in circles*],

---

[17] A differentiation between interviews and news reports in which speakers are sitting versus speakers who are standing is made to uncover whether the smaller gesture space that is available when sitting has an influence on speakers' gesture rate. Furthermore, "sitting" in most cases corresponds to "indoor settings", whereas in outdoor reports and interview, people are usually standing upright.

literal uses of [N *spin around*], i.e. uses that refer to real, physical motion, incite speakers to gesture more than metaphorical uses do[18] (Table 5).

	[N *spin around*]	
	gesture	no gesture
absolute	110	42
relative	72.37%	27.63%

Table 4: Occurrences of [N *spin around*], instantiated with or without gesture

	literal		metaphorical		ambiguous	
	gesture	no gesture	gesture	no gesture	gesture	no gesture
absolute	105	36	4	5	1	1
group internal	**74.47%**	25.33%	44.44%	**55.56%**	50%	50%

Table 5: Frequency distribution according to semantic uses

With respect to the genre distribution, table 6 shows that two genres are predominant in the selection of 152 examples: weather reports and talk/late night shows. Both genres also show high gesture rates but they are only moderately higher than expected[19].

genre		relative %in selection	circular gestures	quotient
talk show/late night talk		33.55%	34.55%	1.03
interview	sitting	2.63%	2.72%	1.03
	standing	5.26%	3.64%	0.69
news report/news show	sitting	7.23%	6.36%	0.88
	standing	5.26%	5.45%	1.04
weather report		44.07%	44.55%	1.01
other (speech, trial,…)		1.97%	2.73%	1.39

Table 6: Genre and gesture distribution for [N *spin around*]

---

[18] The absolute number of occurrences for metaphorical uses in general is, however, very low. Therefore no level of statistical significance can be given.

[19] One can thus conclude that the overall frequency distribution given in table 4 is not strongly biased by the strong representation of these two genres in the corpus

Regarding gesture shape, [N *spin around*] is used with circular hand movements that are very similar to those observed with [V(motion) *in circles*][20]. However, the semantics of those gestures is quite different. With [V(motion) *in circles*], it is always Path, as encoded in the satellite, and only in some cases (depending on the verb used) Manner that is depicted gesturally, whereas circular motion is both a Path and a Manner aspect of "spin around" used as a verb.

3.3. [zigzag]

Also [zigzag][21] fits well into the picture, but the absolute number of occurrences in Red Hen is very low (35). In 71.43% of cases, the zigzag path referred to verbally is co-depicted gesturally, i.e. one or both hands are moved in a zigzag pattern. Again, as with the previous two constructions, literal zigzag-motion is much more often depicted[22] by iconic hand motion than is "metaphorical motion"[23].

	[*zigzag*]	
	gesture	no gesture
absolute	25	10
relative	71.43%	28.57%

Table 7: gesture/no-gesture distribution for [zigzag]

	literal		**metaphorical**	
	gesture	no gesture	gesture	no gesture
absolute	15	8	10	2
group internal	**65.22%**	34.78%	83.33%	**16.67%**

Table 8: gesture/no-gesture distribution for [zigzag] according to semantic uses

---

[20] In 65% of cases, the circular motion is carried out by an extended index finger gesture and in 19%, one observes a one-handed palm open-gesture. In 9 % the palm open-gesture is carried out by both hands while in only 7 cases, we see a gesture where the index fingers of both hands are extended and orbiting each other, i.e. one index finger is circulating the other. One finger is pointing down, the other up.

[21] I included all examples in which "zigzag" is used. Syntactically, it can be a noun (11 cases), a verb (1), an adjective (20), or an adverb (3).

[22] The low number of cases, however, again does not allow for statistical testing.

[23] Such as instantiated in exchange (2) from the late night show with Craig Ferguson: "Don Rickles: by the way; you're doing so well; I'm very happy for you; really am; he's really great; ((cheering and applause from the audience)); Craig Ferguson: see; you always do the zigzag; you do like that; and then you go like that".

3.4. *[all the way from X PREP Y]*

The last construction under multimodal scrutiny here is *[all the way from* X PREP Y]. Table 9 merges aspects of use with absolute and relative gesture-occurrence.

Use	[*all the way from X PREP Y*]	
	gesture	no gesture
area	110 (85.94%)	18
distance	40 (85.11%)	7
development/time span	5 (55.56%)	4
list/spectrum	5	10 (66.67%)
Total	**160 (80.4%)**	**39 (19.6%)**

Table 9: Frequency table for multimodal instantiations of [*all the way from* Y PREP Y]

With 80.4%, the general co-occurrence rate is the highest observed so far. Similar to the previously analyzed constructions, the use of co-speech gesture is, again, strongly influenced by the semantic uses of the construction. When speakers use the construction to refer to or to delineate a specific physical area or distance, they depict this area or distance also with gestural means in approximately 85% of cases. However, when they refer to a development in time or a spectrum of abstract entities that stretches "all the way from X PREP Y", they use far less hand gestures to co-depict this range[24].

Regarding form and function, the variation of the observed gestures is bigger than with the previous three constructions. By far the biggest portion of the data (65%) comes from weather reports. There, the co-speech gestures used with [*all the way from* X PREP Y] usually have both a deictic and an iconic function. That is, the weatherwoman talks about a given weather phenomenon that either affects a certain area (1st use) or stretches over a certain distance (2nd use). Thereby she marks the beginning of that area (either on the weather chart or in the air) by performing a pointing gesture (usually by one hand, which is subsequently held) and then performs a motion gesture with the other hand and marks the endpoint of the area/distance by a second deictic gesture (ex. 3, Fig. 5). Alternatively, speakers use the same hand to perform consecutive strokes with movement phases in between (ex. 4, Fig. 6).

---

[24] The numbers for both metaphorical uses are, however, too small to draw any definite conclusions.
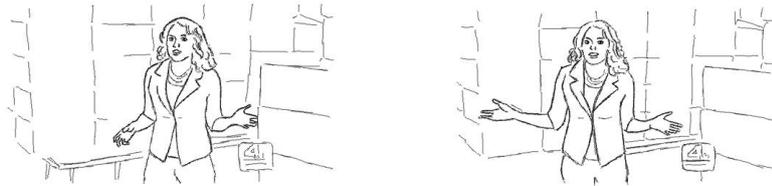
Fig. 5: Two-deictic-stroke-gesture for ex. (3): "the food is delivered frozen each day **all the way from Long Beach to Lancaster**" (KNBC 4 News at Noon, December 25, 2012). The distance in between the two extended hands metaphorically maps onto the distance between Long Beach and Lancaster.

Fig.6: Deictic multiple-stroke gesture for ex. (4): "because ah the storm system is up to the north and gonna affect the folks up **all the way from da the the the Seattle area down through Oregon into the Bay area.**" (Red Hen, KCBS CBS 2 News at 5, December 21, 2012)

The strokes have deictic function as they indicate/virtually place end points or significant points within an area on a map/ in gesture space (McNeill 1992). The motion phase in between is iconic as it depicts the area's/distance's largeness that is expressed by "all the way" (with "all" usually being stressed[25]) by stretching over a large part of the gesture space.

The form of the gestures observed with the "development in time"-use (ex. 5) and the "list/spectrum"-use (ex. 6) are very similar to those described for the other two uses. However, their function is not deictic and/or iconic. Instead, they are motivated by spatial metaphoric mappings (Mittelberg 2006, Cienki & Müller 2008).

Ex. (5): "…to make sure that every single child in this nation has a world class education **all the way from preschool to college."** (Red Hen, MSNBC News Live, September 4, 2012)

---

[25] The relationship between gestural instantiation and prosody has, however, not been studied systematically.

all the way……………………from preschool

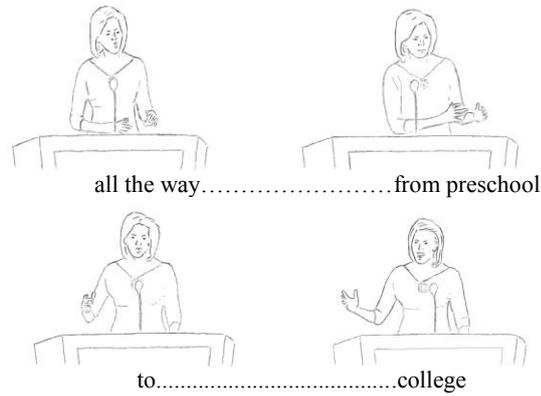to.......................................college

Fig. 3: Example of a (development in) time use and a gesture co-depicting the metaphorical mapping of time conceptualized as movement in space

Ex. (6): "maybe I didn't fully appreciate where we were going, but there was a whole system going on, **all the way from the borrower of the mortgage, all the way through to the investor"** (Red Hen, KOCE Frontline, April 24, 2012)
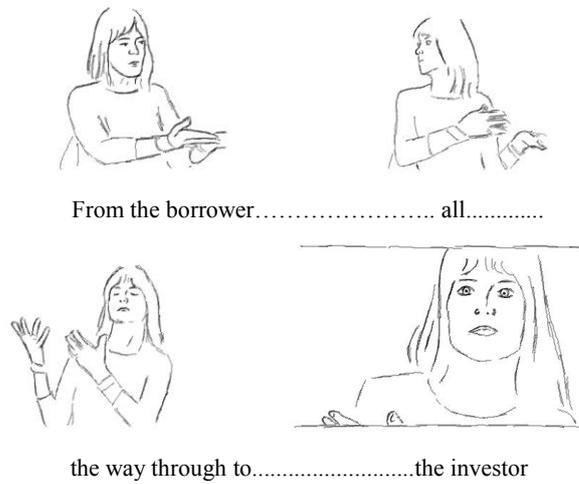

From the borrower………………….. all.............

the way through to...........................the investor

Fig. 4: Example of metaphorical gesture observed with a spectrum reading of [*all the way from* X PREP Y]

4. Conclusions

This paper gave a very compact overview of recurrent hand gestures accompanying the constructions [V(motion) *in circles*], [*zigzag*], [N *spin around*], and [*all the way from* X PREP Y]. All four constructions show strong associations with recurrent forms of gestures, ranging from 60% to 80%. Specific semantic uses (literal motion with the first three constructions, area/distance-uses with [*all the way from* X PREP Y]) incite speakers to gesturally depict a specific motion aspect or use a deictic/iconic gesture even in 65% to 85% of cases. Applying the Construction Grammar criterion of "recurrence", the paper therefore shows that the use of co-verbal gesture is indeed a recurrent usage aspect of these constructions. Hence, it is suggested that these constructions may not only have verbal but also gestural structure.

The quantitative analyses, however, suffer from a fundamental qualitative problem: Construction Grammar only quite vaguely posits that constructions are entrenched form-meaning pairings if they reoccur with "sufficient frequency" (Goldberg 2006:5). What frequencies rates are to be regarded as sufficient, i.e. recurrent enough, to serve as proofs of mental entrenchment, however, is unclear. At first sight, the numbers presented here may even seem deceiving because the gesture-construction co-occurrence rate never is 100%. Such a 100% match would indeed be an irrefutable proof of the gestures' entrenchment. However, this should not automatically lead to the conclusion that a 60 or 80%-occurrence rate for gestures that strongly resemble each other with respect to their form, their temporal placement and their function is no sign of entrenchment. Studies from Interactional Construction Grammar (e.g. Auer & Pfänder 2011) have shown that speakers do not simply instantiate constructions as pre-packaged wholes but rather orient towards constructions in conversation, i.e. they adapt their form and meaning according to in situ conversational requirements. This observation might also hold for the gestural components of constructions. After all, from a usage-based perspective, it seems plausible that when many different speakers do very similar things, i.e. use constructions together with similar hand gestures, in very different circumstances and across conversational genres, this may be due to at least some degree of entrenchment and conventionalization. Nonetheless, whether gestural components are instantiated in usage or not may not only be a matter of entrenchment but also of local, conversational constraints.

Therefore, one of the biggest challenges for Construction Grammar in general and a multimodal extension thereof in particular, is to make its core criterion of "sufficient frequency" operational (Traugott & Trousdale 2013). This entails the need of a baseline for quantitative comparison of specific construction-gesture co-occurrence rates. In other words, to the best of my knowledge, it is unknown

how often, on average, just any construction is used together with hand gestures. This knowledge, however, is a prerequisite for knowing whether a given observed co-occurrence rate is an indication of specific constructional entrenchment. Despite this lack of a means for quantitative comparison, the frequencies observed for the four constructions under scrutiny (60-80%) in this paper seem to be high enough to at least put into doubt the Construction Grammar perspective on constructions as purely verbal form-meaning pairings.

5. References

Auer, P. & Pfänder, S., eds, (2011) *Constructions: emerging and emergent*. Mouton de Gruyter, Berlin.

Andrén, Mats (2010) *Children's Gestures from 18 to 30 months*. Lund University: Centre for Languages and Literature, Lund.

Cienki, A. & Müller, C., eds, (2008) *Metaphor and gesture.* John Benjamins, Amsterdam/Philadelphia.

Goldberg, A. (1995) *Constructions: a construction grammar approach to argument structure.* University of Chicago Press, Chicago.

Goldberg, A. (2006) *Constructions at work*. Oxford University Press, Oxford.

Hickmann, M., Hendriks, H. & M. Gullberg (2011) How to talk and gesture about motion in French and English: cross-linguistic and developmental perspectives. In *Langage, Interaction et Acquisition 2, 1. Grammaticalization and language acquisition: nouns and verbs across languages*, 129-156.

Johansson Falck, M. & Gibbs, R. (2011) Embodied motivations for metaphorical meanings. *Cognitive Linguistics* 23.2, 251-272.

Kendon, A. (2004) *Gesture: Visible Action as Utterance.* Cambridge University Press, Cambridge.

Kita, S. & Özyürek, A. (2003) What does cross-linguistic variation in semantic coordination of speech and gesture reveal? Evidence for an interface representation of spatial thinking and speaking. *Journal of Memory and Language* 48.1, 16-32.

Ladewig, S. (2011) Putting the cyclic gesture on a cognitive basis. *Cognitextes* 6. http://cognitextes.revues.org/406

Langacker, R.W. (2001) Discourse in Cognitive Grammar. *Cognitive Linguistics* 12, 143-188.

Langacker, R.W. (2008) Metaphoric gesture and cognitive linguistics. In A. Cienki & C. Müller, eds, *Metaphor and gesture*. John Benjamins, Amsterdam, 249-251.

McNeill, D. (1992) *Hand and Mind: What Gestures Reveal about Thought.* University of Chicago Press, Chicago.

McNeill, D. & Duncan, S. (2000) Growth points in thinking-for-speaking. In D. McNeill, ed, *Language and Gesture*. Cambridge University Press, Cambridge, 141-161.

Mittelberg, I. (2006) *Metaphor and metonymy in language and gesture: Discourse evidence for multimodal models of grammar*. Dissertation, Cornell University.

Nicoladis, E. & Brisard, F. (2002) Encoding motion in gestures and speech: are there differences in bilingual children's French and English. In E. Clark, ed, *Space in*

language: location, motion, path and manner: proceedings of the 31st Stanford Child Language Research Forum. Center for the Study of Language and Information, Stanford, 60-68.

Praggelaz Group (2007) MIP: A Method for Identifying Metaphorically Used Words in Discourse. *Metaphor and Symbol* 22.1, 1-39.

Richardson, D. & Matlock, T. (2007) The integration of figurative language and static depictions: An eye-movement study of fictive motion. *Cognition* 102, 129-138.

Schoonjans, S. (in preparation). *Modalpartikeln als multimodale Konstruktionen. Eine korpusbasierte Kookkurrenzanalyse von Modalpartikeln und Gestik im Deutschen*. PhD thesis. University of Leuven.

Steen, F. & Turner, M. (2013) Multimodal Construction Grammar. In M. Borkent, B. Danvygier & J. Hinnell, eds, *Language and the Creative Mind*. CSLI Publications, Stanford.

Streeck, J. (2008) Metaphor and gesture. A view from the microanalysis of interaction. In A. Cienki & C. Müller, eds, *Metaphor and gesture*. John Benjamins, Amsterdam, 259-264.

Talmy, L. (1985) Lexicalization patterns: Semantic structure in lexical forms. In T. Shopen, ed, *Language typology and semantic description. Vol. 3: Grammatical categories and the lexicon*. Cambridge University Press, Cambridge, 36-149.

Talmy, L. (2000) *Toward a Cognitive Semantics*. MIT Press, Cambridge, MA.

Traugott, E. & G. Trousdale (2013) *Constructionalization and Constructional Changes*. Oxford University Press, Oxford.

Wilson, N.L. & Gibbs, R.W., (2007) Real and imagined body movement primes metaphor comprehension. *Cognitive Science*, 31, 721-731.

Zima, E. (submitted) *'He ran all the way from Montreal to Atlanta, that is when he stopped running and began walking'. A multimodal view on motion constructions*.

Zima, E. (in press) Gibt es multimodale Konstruktionen? Eine Studie zu englischen und deutschen Bewegungskonstruktionen. *Gesprächsforschung Online*.