The influence of native language word order and cognitive biases in pictorial event representations

Vastenius, Anu; van de Weijer, Joost; Zlatev, Jordan

Published in:
Cognitive Semiotics

DOI:
10.1515/cogsem-2016-0004

2016

Document Version:
Publisher's PDF, also known as Version of record

Anu Vastenius*, Joost van de Weijer and Jordan Zlatev

The influence of native language word order and cognitive biases in pictorial event representations

DOI 10.1515/cogsem-2016-0004

Abstract: Ever since Goldin-Meadow, S., C. Mylander, W. C. So, and A. Özyürek. 2008. The natural order of events: How speakers of different languages represent events nonverbally. PNAS 105: 9163–9168. proposed that there is a preferred order in sequential non-verbal event representations (Actor > Patient > Act), apparently independent of the default word order in one's native language, the topic has been the focus of much cognitive-semiotic research. After providing a partial review of the field, we describe an empirical study investigating the order of pictorial representations of motion events using a design that emphasized the linearity of the representations to a greater extent than Goldin-Meadow et al. (2008). Speakers of Swedish (default word order: Actor > Act > Patient, or SVO) and speakers of Kurdish (default word order: Actor > Patient > Act, or SOV) participated in the study. Unlike earlier studies, we found an effect of native language word order. The Swedish speakers preferred to place the Patient picture after the Act picture, especially after first describing the stimuli verbally. In contrast, the Kurdish speakers preferred Act after Patient both with and without verbalization. The results of the study suggest that any cognitive or communicative biases for particular constituent order in non-verbal representations are likely to be modulated by linguistic word order, at least in populations reliant on written language in their daily lives.

Keywords: constituent order, word order, non-verbal representations, thinking for speaking, motion events, linguistic relativity

1 Introduction

How do people represent events using non-verbal semiotic vehicles such as gestures or pictures? Is the sequential order of their non-verbal representations influenced by the basic word order of their native languages? Or is there a

***Corresponding author: Anu Vastenius,** Centre for Languages and Literature, Lund University, Box 201, 221 00 Lund, E-mail: anu.vastenius@semiotik.lu.se
Joost van de Weijer, Jordan Zlatev, Centre for Languages and Literature, Lund University, Box 201, 221 00 Lund

particular order, one based on general cognitive-semiotic biases, that is generally preferred? Such questions have been addressed in a number of studies that have, interestingly, provided rather different answers (e. g., Kita and Özyürek 2003; Goldin-Meadow et al. 2008; Christensen, Riccardo, and Kristian 2016). In one of the most influential studies, Goldin-Meadow et al. (2008) tested how speakers of languages with different basic word orders – Subject-Verb-Object (SVO) or Subject-Object-Verb (SOV) – represented events non-verbally with pictures and gestures. The results suggested that all groups predominantly used the order Actor-Patient-Act (ArPA) in their nonverbal representations, irrespective of task and native language word order, leading to the conclusion that "there appears to be a natural order that humans, regardless of the language they speak, use when asked to represent events non-verbally" (Goldin-Meadow et al. 2008: 9167). Similar experiments with only gestures have been conducted since then (Langus and Nespor 2010; Meir et al. 2010; Schouwstra 2012; Hall, Mayberry, and Ferreira 2013; Christensen and Tylén 2013; Christensen, Riccardo, and Kristian 2016), showing a more complex picture. Furthermore, the pictorial task used by Goldin-Meadow et al. (2008) raises some methodological issues. Curiously, the order of spontaneous sequential event representations using pictures remains underexplored.

To throw light on these issues, we here report on a study of event representations using pictures that are practically identical with those used by Goldin-Meadow et al. (2008), with one key difference: rather than being printed on transparencies that were to be stacked on top of one another, they were printed on cardboard cards, to be placed in a line. To test for an effect of linguistic word order, one group of speakers had Kurdish, with Subject-Object-Verb (SOV) as predominant word order, while the other group had Swedish, with Subject-Verb-Object (SVO) as default word order. To exclude the influence of the multilingual environment in Sweden where most speakers with native SOV languages have at least basic knowledge of Swedish, the study with Kurdish speakers was conducted in a location where it is the dominant language in the society, Iraqi Kurdistan. The methodology of the study is described in detail in Section 4, while Section 5 presents our major findings, and Section 6 discusses its implications.

Prior to that, we review the field and its theoretical relevance for cognitive semiotics in Section 2, discussing both the issue of the impact of language on cognition and the various factors that have been suggested to influence constituent orders in non-verbal representations. In Section 3, we introduce the topic of word order as a typological feature in language and provide concrete examples of the two languages addressed in the study, Kurdish and Swedish. Finally, conclusions and suggestions for future research are given in Section 7.

2 Theoretical background: Linguistic relativity and non-verbal event representations

At least two active fields of current research form the backdrop to our investigation: the "neo-Whorfian" trend of demonstrating (often subtle) effects of native language on performance in non-verbal cognitive tasks as well as the field of non-verbal sequential event representations, mentioned in the introduction. The two have been seldom considered in tandem, which is why we here provide brief reviews of each, with a focus on the latter.

2.1 Linguistic influence on thinking

Does the language we speak affect the way we think? This question has been addressed many times since von Humboldt argued that languages carry their own "world views" (Humboldt, [1836], 1988). A century later, Whorf (1956) famously introduced "the principle of linguistic relativity:"

> Users of markedly different grammars are pointed by their grammars towards different types of observations and different evaluations of externally similar acts of observation, and hence are not equivalent as observers but must arrive at somewhat different views of the world (ibid.: 221).

In the last decades of the previous century. such ideas fell under the attack of universalist and modularist theories in cognitive science (Fodor 1983; Pinker 1994): "The famous Sapir-Whorf hypothesis [...] that differences among languages cause differences in thoughts of their speakers [...] is wrong, all wrong (Pinker 1994: 57). Pinker (1994) argued that language and thought are separate from one another based on the existence of thinking in deaf isolates, aphasics, and pre-verbal infants. However, the fact that not all thinking needs language does not exclude the possibility that language may still affect thinking, and indeed evidence for this has been reported (e. g. Levinson 2003; Casasanto 2008). Research in linguistic typology has shown that linguistic structures indeed differ substantially in the world's languages (Givón 1979; Gell-Mann 2011; Dixon 2012; Dryer 2013). It is thus possible that the language-specific patterns that we use in communication can have some impact on the way we think.

The evidence for linguistic influence on thought comes from diverse sources and methodologies. For example, it has been shown that spatial metaphors for time are prevalent in the world's languages. Still, which spatial concepts correspond to which aspects of time varies. In European languages, we talk about the

future being in front of us, but in Aymara the future is spoken of as the space behind us, and these differences are also reflected in co-verbal gestures (Nunez and Sweetser 2006). In English, linear spatial expressions are used to describe time such as *long* vs. *short* time, but in Greek the expressions denote volume: *megálos* ('large') vs. *mikrós* ('small') and *polí* vs. *lígo* ('much' vs. 'little'). In an experiment with English and Greek speakers, participants performed according to the metaphors in their respective languages in a non-linguistic task involving the estimation of duration (Casasanto et al. 2004). Notably, however, the English speakers could be induced to behave like the Greeks after a relatively brief period of exposure to volume-based metaphors for time.

In an attempt to develop a more generally acceptable theory of linguistic relativity, Slobin (1996) presented the idea of "thinking for speaking," proposing that language has (at least) an impact on the kind of thinking that goes on when actually speaking.

> Whatever effects grammar may or may not have outside of the act of speaking, the sort of mental activity that goes on when formulating utterances, is not trivial or obvious, and deserves our attention. [...] In the evanescent time frame of constructing utterances in a discourse, one fits one´s thought into available linguistic frames, and picks those characteristics of objects and events that (a) fit some conceptualization of the event, and (b) are readily encodable in the language. (Slobin 1996: 76)

This hypothesis has the implication that "languages differ from one another not only in the presence or absence of a grammatical category, but also in the way in which they allocate grammatical resources to common semantic domains" (Slobin 1996: 83). For example, languages have been argued to differ in their preferred lexicalization patterns for expressing motion events. English and German have been analysed as *satellite-framed languages* (expressing Path/Direction with particles, and having many verbs for expressing manner-of-motion, like *stagger*, *slither*, and *creep*), while French and Hebrew are *verb-framed languages* (Talmy 2000), typically expressing Path/Direction in verbs like *entre* and *sortir* ('exit'). When children acquiring these languages were asked to describe scenes in a children´s book, speakers of Spanish and Hebrew used significantly more descriptions that had no locative elaboration, and tended to use a Path/Direction verb when describing a downward motion. In contrast, the speakers of English and Germen tended to encode the manner-of-motion more often, and in greater detail. To some degree, this may be taken as supporting Slobin's hypothesis.

One of the best-documented studies for linguistic influence on thinking involves different strategies for specifying the location of an object in space (Levinson 2003; Majid et al. 2004). In a language with "relative" frame of

reference, the position of an object is typically described as being to the *left* or *right* of a landmark. A speaker of a language where the "intrinsic" frame is widely employed would tend to describe the same spatial relation with reference to some feature of the landmark, for instance, as *on the back side of X*. Finally, in a language with an "absolute" frame, the position would be described as being *to the North of* or *to the South of* the landmark object. In languages with absolute frame such as Tzeltal and Guugu Yimithirr, speakers use this frame to describe everyday small-scale locations. For instance, a speaker of Guugu Yimithirr could provide a spatial description corresponding to *The fork is to the North of the spoon* (Majid et al. 2004). Importantly, studies have demonstrated performance in non-linguistic tasks in accordance with the dominant frame of reference of one's language. For example, Majid et al. (2004) describe an experiment where participants were asked to place an object in the "same" position after rotating 180°. Dutch participants solved this according to their dominant relative frame, while speakers of Tzeltal used the absolute frame. These and other similar findings lead to the conclusion that while linguistic determinism may be unfeasible, a degree of linguistic influence on thinking is possible, and even likely (Zlatev and Blomberg 2015).

Features of one's language have also been shown to influence one's use of gestures. Studying speakers of English and Spanish, McNeill (2000) showed differences in their co-speech gestures that could be co-related with typological features of the respective language. Kita and Özyürek (2003) investigated how clausal packaging of semantic structure may affect spontaneous gesturing by eliciting narrative descriptions of cartoon animations. Speakers of English, Turkish, and Japanese provided different iconic gestures when describing the same motion events, mostly depending on whether the descriptions involved a single verb (English) or two separate ones (Turkish, Japanese). On this basis, the authors proposed that "if language-specific spatial representation is repeatedly generated for speaking, then it can become part of habitual non-linguistic thought about space, that is, *the default way of thinking about space even outside the context of speaking*" (ibid.: 27). In other words, linguistic structuring of event representations could potentially affect how events are imagined (in thought) and (re-)represented in other semiotic media, like gestures and pictures.

2.2 Constituent orders in non-verbal representations

Despite the findings summarized above, the predominant view in the field of sequential non-linguistic event representations is that there exists a basic constituent order that is independent of the structure of one's native language.

This conclusion is largely due to the study of Goldin-Meadow et al. (2008), mentioned in the introduction, in which speakers of different languages were asked to represent events using silent gestures and pictures. The participants in the study were forty adults (ten per language): speakers of English and Spanish, both with the predominant word order SVO, Turkish with SOV order, and Chinese with a mixed order. They were shown video clips displaying 36 different motion events (see Section 4, Table 2).

In the first experiment, participants were asked to describe all the clips first in their native language. Then the clips were shown again, and their content was to be represented by means of gestures only. In the second experiment, transparencies with the pictures of each "actant" (see Table 2 in Section 4.1) were to be stacked on top of each other so as to form a composite picture of the event. The order in which the transparencies were picked did not affect the final product: the image would always end up the same. The order of constituents used in the gesture task and the order of picking the transparencies were then coded. Importantly, the participants were not found to display the orders of their spoken languages. Rather, the order of the constituents both in the gesture and in the transparency task was predominantly Actor-Patient-Act (ArPA). Based on these findings, the authors proposed that Actor (Ar) and Patient (P) are chosen before Act (A) as they are cognitively less demanding, in being less relational. In addition, there is a "particularly close cognitive tie between objects and actions, which would link Patient to Act, resulting in an ArPA order" which "may reflect a natural disposition that humans exploit not only when asked to represent events nonverbally, but also when creating language anew" (Goldin-Meadow et al. 2008: 9166–9167).

The corresponding word order, Subject-Object-Verb (SOV), has been suggested to be the original order in human languages in a number of independent studies. The direction of syntactic change is mostly from SOV to SVO (Givón 1979; Gell-Mann and Ruhlen 2011). For example, Proto-Indo-European was probably SOV, while its modern decedents most often have SVO-order. A functional motivation for the change commonly mentioned is that it is easier to keep Subject and Object apart cognitively if they are separated structurally, especially in grammatically more complex languages (Langus and Nespor 2010; Gibson et al. 2013), especially if they lack a grammatical case to help keep the grammatical functions apart. In fact, most SOV languages do have case systems (Dryer 2013). Another piece of evidence that SOV is in some sense the original order comes from spontaneously emerging languages, such as Al-Sayyid Bedouin Sign Language (Sandler 2010). Furthermore, "home-signers" (individuals deaf from the birth, who have not had a possibility to learn a signed language), tend to develop gesture-based communication systems with preferences for SOV order (Goldin-Meadow et al. 2008).

Can we therefore conclude that the Actor-Patient-Act order is in general cognitively basic and therefore most often used when representing events non-verbally? A number of different findings stand in the way of a simple "yes" to this question. To begin with, in the study of Goldin-Meadow et al. (2008), half of the video-clips contained a spatial endpoint, or Landmark (LM). In representing these clips with gestures, speakers of all four languages (Turkish, Chinese, Spanish and English) tended to place the LM at the end of the sequence. On the other hand, in the picture transparency experiment there was a clear bias for placing the LM first, which at least in Turkish corresponds to preferred word order. This indicates that different experimental designs may result in different outcomes.

Other researchers have suggested alternative hypotheses regarding the factors influencing constituent order in non-verbal representations. Langus and Nespor (2010) proposed that the two word orders SVO and SOV originate from different cognitive systems: in "simple" communication with direct interaction between the sensory-motor and the conceptual system, SOV may be the preferred order, but in "the computational system of grammar (syntax)" (Langus and Nespor 2010: 292), SVO may be favoured. Thus, word order differences across languages "emerge from the struggle between individual cognitive systems trying to impose their preferred structure on human language" (Langus and Nespor 2010:: 291). In addition to what was mentioned earlier, support for this may be found in pidgin languages, which most often have SOV order – but with increased grammaticalization, they often adopt SVO order. Langus and Nespor (2010) performed four studies to investigate these claims: a task with (a) gestural descriptions of simple scenarios, (b) gestural descriptions of complex scenarios, (c) gestural comprehension, and (d) comprehension of flattened speech, i. e., artificially flattened prosodic strings. Speakers of two languages, Italian (SVO) and Turkish (SOV), were tested. In simple gestural tasks, ArPA was the preferred order of constituents, but in the more complex ones, ArAP was the favoured order. In the gestural comprehension task, SOV elicited the fastest reactions in both language groups. When the performance of the two language groups was compared in relation to the order of Object and Verb, both groups had faster responses when Object preceded Verb, supporting the idea that "improvised gesture production, as well as comprehension, is not mediated by the computational system" (Langus and Nespor 2010: 308). However, both groups had shorter reaction times in task (d) when the order of their respective native language was used in the flattened word strings.

In another recent study, Hall et al. (2013) suggest that SOV is not optimal for describing an event in which Actor and Patient can be interchanged. They point out that case marking, common in SOV languages as mentioned earlier, is

especially important when describing *reversible* events, i. e. events where both Subject and Object refer to animate beings, and each could thus potentially take the role of Agent. The authors elicited gestural descriptions of video clips showing reversible and non-reversible events. The results indicated that ArPA was the most common order among non-reversible events, whereas among reversible events it was ArAP in all conditions, leading to the conclusion that "semantic properties of the event are among the factors that determine which order is favoured" (Hall et al. 2013: 15).

Meir et al. (2010) proposed a similar hypothesis: different kinds of clauses give rise to different communicative challenges. Transitive clauses can either be *canonical,* where the Subject/Actor is animate and the Object/Patient inanimate, or *reversible,* where both arguments are animate, and a grammatical strategy is needed to mark the Actor. The authors tested if there was a difference in constituent order for each kind of clause when representing events by gestures. Thirty-three native speakers of Hebrew (a SVO language) gestured events presented in video clips; in clauses with an inanimate object ArPA was more common than ArAP (65 % vs. 31 %.) In reversible clauses, the reverse pattern was found: ArAP 64 % and ArPA in 31 %. In addition, participants invented additional devices, for instance showing initially two fingers to indicate the number of agents. The authors concluded that SOV might be the basic word order, and the others have developed later as a response to processing efficiency and communicative demands.

Another kind of factor was investigated by Schouwstra (2012), who hypothesised that so-called intensional events (1a) will tend to elicit ArAP rather than ArPA orders, in contrast to extensional motion events like (1b).

(1) a. The princess wants an apple.
b. The princess throws an apple.

The ontological status of the Agent is not the same as that of the Patient in (1a): the princess has to exist for the sentence to be true, while the desired apple does not necessarily exist. The verbs classified as intensional were "classical" intensional verbs *(search*), psychological verbs (*dream of, think of),* perception verbs *(hear, see),* and creation verbs *(build, draw, knit, sculpt)*. In a first experiment, sixteen speakers of Dutch, a SVO-language, were shown pictures of 20 motion events and 20 intensional events, and asked to communicate their content to the experimenter through gestures. The results showed that there was a significant correlation between the picture type and the order of constituents: in motion events the participants predominantly used ArPA order, but in intensional events, the most usual gestural order was ArAP. In a second experiment,

short movie clips were presented showing an actor who gestured simple events that could be interpreted either as motion events or intensional events, for example (2a) and (2b).

(2) a. A witch climbs a house.
b. A witch builds a house.

Both events were gestured in exactly the same way, but the order of the Act and the Patient was switched. The participants were asked after each clip to choose one of two pictures that corresponding to an intensional and a motion event, respectively. The results showed that for ArPA strings of gestures, a motion event interpretation was preferred, while for ArAP orders an intensional event reading was most often chosen.

Additionally, 19 speakers of Turkish, a SOV-language, participated in an experiment conducted by Schouwstra and de Swart (2014) with the design used in the first experiment of Schouwstra described above. Similar results were obtained: just like the speakers of Dutch, the speakers of Turkish used predominantly ArPA for motion events and ArAP for intensional events, while there was no significant difference between the language groups. The authors conclude that:

> [...] the fine-grained semantic properties of the message influence sequencing of information, and that this is a dynamic process that happens 'on the fly'. This sheds a novel light on the possible mechanisms at work in emerging languages. There is no pre-set basic word order, but a range of possible linearization options, and choices are driven by different factors. (ibid.: 435)

In another study, a similar approach was adapted, though using a different conceptual apparatus. Christensen and Tylén (2013) hypothesized a different order in gesturing based on the difference between *object manipulation events* and *construction events*. (3a) is an example of a representation of a manipulation event, and (3b) a representation of a construction event.

(3) a. A ballerina is throwing a paper plane.
b. A ballerina is painting a paper plane (on canvas).

The authors designed a game in which pairs of participants communicated the two kinds of transitive events in gestures. Two sets of pictures were created, so that each Actor and Patient appeared in both manipulation and construction events, which forced participants to disambiguate the elements, and to include them all. The results confirmed the prediction: in manipulation events, the dominant order turned out to be ArPA and in construction events ArAP.

Christensen and Tylén (2013) argued that motivational sources can be revealed in conceptualization of events by gestures, and that the gesture strings "iconically reflect the inherent logical and sequential structure of the perceived events" (Christensen and Tylén 2013: 1). According to the authors, the conceptual structure in gesture is influenced by external, real-world factors and thereby is not restricted to the cognitive domain. We may note that the two categories of events used by the authors can be categorized as sub-types within the event types of Schouwstra (2012): motion events and intensional events, respectively. In "manipulation events," objects exist prior to the action, but in "construction events," they are the consequence of the action and thus have a different sequential and logical structure.

To sum up, these studies reveal the influence of the following factors on the order of non-verbal representations:

- The reversibility in transitive events (with animate vs. inanimate Patient)
- The ontological status of the Patient (intensional vs. extensional event, manipulation vs. construction event)
- The complexity of the message (simple vs. complex)

It is intriguing that when the designs include speakers of languages with different word orders (Turkish has invariably been used as representative of the SOV category), no effect on gesture order preferences was found. Still, it seems that native language has not been sufficiently studied as a possible determining factor. After all, its influence on cognition and gestures has been shown using other experimental paradigms, as summarized in Section 2.1.

Even more underexplored is the use of sequential pictorial event representations. In the study of Goldin-Meadow et al. (2008), these were, to remind, printed on transparencies which were to be stacked atop one another to form a composite representation. However, this does not reflect a typical way of using pictorial representations by either adults or children. A more common manner of arranging pictures representing different participants in the event would be to use a given *sequential* order (left-to-right, right-to-left, top-down) as in comic books. It has not been tested if an image of the Act (represented, e. g., as an arrow) would still tend to be placed at the end of such a sequence or if it would it be positioned more centrally, especially by speakers of SVO languages. Furthermore, if participants were asked to describe the events first verbally, and only then non-verbally, one may expect a stronger effect of the respective linguistic features, per Slobin's "thinking-for-speaking" hypothesis (see Section 2.1). Thus, we can formulate the following research questions:

- Does the basic word order of one's native language have an impact on the constituent order used in a pictorial representation of a motion event?

- Is there a stronger influence of native language word order if the event is described verbally prior to representing it pictorially?
- Will the manner of presenting and placing pictures give rise to different results: on transparencies + one over the other vs. on picture cards + sequentially?

These questions are addressed in the experiment described in Section 4. Prior to this, however, we provide some linguistic description of the two languages involved.

3 Word order in Swedish and Kurdish

In this section, we consider a basic feature of constituent orders in language: the alignment between grammatical functions and semantic roles, and describe how Swedish and Kurdish differ in this respect. According to current estimates, SOV is the most common word order worldwide. The second most common is SVO, and languages with no dominant word order follow (see Table 1, after Dryer 2013). While there is considerable intra-linguistic variation, most languages have a dominant, unmarked word order. Deviations from this tend to express marked constructions, such as passives and topicalizations (Gell-Mann and Ruhlen 2011).

Table 1: Order of subject, object and verb in a large, balanced sample of the world's languages (Dryer 2013).

Dominant order	SOV	SVO	No dominant order	VSO	VOS	OVS	OSV	Total
Number of languages	565	488	189	95	25	11	4	1377

Constituent order in language can be formulated in two ways: by using (a) grammatical functions and (b) semantic roles. Grammatical functions such as *subject*, *direct object* and *indirect object* signify the grammatical role played by noun phrases in a clause. Intransitive sentences contain a subject only, while mono-transitive and di-transitive sentences contain one or two objects respectively. Semantic roles are rather the types of arguments in the predication expressed by the clause, such as Actor, Patient, Recipient, and Goal (Saeed 2009). There is no definitive consensus in the field as to what the roles are or how they should be distinguished. In the present study, three semantic roles are utilised: Actor (Ar), Patient (P), understood as prototype-based (see below) and

Landmark (LM) in addition to the main predicate (A). There is a general pattern of alignment, although with exceptions, between the semantic roles and the grammatical functions (Givón 1990). In an *accusative* language such as English, the Actor is often the grammatical subject, the Patient is the direct object, and the Goal or Recipient (here called Landmark) is usually expressed by the indirect object. Since there are no grammatical functions in non-verbal representations, semantic roles are of greater importance in discussing the order in pictures, and their possible similarity with the constituent order of sentences.

Dowty (1991) proposed that only two kinds of roles are needed for alignment: Proto(typical)-Actor and Proto(typical)-Patient. In our study, Actor and Patient are similarly understood as prototype-based concepts, defined by clusters of entailments. Not all typical entailments need to be present (volition, intention, animacy etc.) for something to be an Actor. The central feature of an Actor is taken to be *causal force*, while the Patient is *affected* by this force. To these two proto-roles, we can add Landmark, also as a prototype-based role, which (together with directional prepositions) designates the destination of a motion event.

The basic word order in most Indo-European languages is SVO, corresponding to ArAP in terms of semantic roles, as is the case in English (4). The default word order in Swedish is similar to English, as shown in (5).

(4) I ate an apple.
S V O (functions)
Ar A P (roles)

(5) Jag äter ett äpple.
I eat.PRES DET.INDEF apple
Ar A P
"I am eating an apple."

In Swedish, when there is an indirect object in the clause with the semantic role of Landmark (Goal), it is usually placed after the verb as in (6). The Landmark (when also serving the role of Recipient) may either precede or follow the Patient, as shown in (7a) and (7b), similar to English.

(6) Vi går till centrum.
We go.PRES to centre
Ar A LM
"We are going to the centre."

(7) a. Han ger en blomma till flicka-n.
He give.PRS DET.INDEF flower to girl-DEF
Ar A P LM
"He gives a flower to the girl."
b. Han ger flicka-n en blomma.
He give.PRS girl.DEF DET.INDEF flower
Ar A LM P
"He gives the girl a flower."

If there is an adverbial, an adverbial clause, or a dependent clause in the beginning of the sentence, the constituent order changes, so that the subject (Actor) follows the verb, as Swedish is a so-called "V2 language" (Josefsson 2004). This, however, does not change the relative order between Ar, P and LM, as shown in example (8a):

(8) a. Ofta ger han en blomma till flickan.
Often give.PRS he DET.INDEF flower to girl.DEF
A Ar P LM
"Often he gives a flower to the girl."

On the other hand, placing the Recipient/LM (8b) or the Patient (8c) in initial position does of course change the relative order of the three semantic roles Ar, P, and LM. Hence, this is something to take into account when looking for correlations between the order of constituents in linguistic descriptions and pictorial representations in our study.

(8) b. Till flickan ger han en blomma.
LM A Ar P
(8) c. En blomma ger han till flickan.
P A Ar LM

Kurdish belongs to the Indo-Iranian branch of Indo-European languages, and more precisely to the West-Iranian group. It is spoken by approximately 30 million speakers in Turkey, Iraq, Iran, and Syria (www.ethnologue.com). The two main dialects are Kurmanji and Sorani, which are not entirely mutually intelligible due to differences at the structural level and in vocabulary (Thackston 2006). The participants in our study were speakers of the Sorani dialect spoken in Suleymania-region in northern Iraq. The basic word

order is SOV, or Actor-Patient-Act, and the Patient needs to precede the Act as shown in (9).[1]

(9) a. Mn seu dakhom.
I apple eat.PRES
Ar P A
"I am eating an apple."
b. *Mn dakhom seu.

The Landmark usually comes after the verb (Haig 2002), as shown in (10a). However, it can also be placed between the Ar and A, as in (10b), or even in initial position, as in (10c).

(10) a. Ema da-rroi-in bo nau shaar.
We PRES go 1.PL to centre
Ar A LM
"We are going to the center."
b. Ema bo nau shaar darroiin.
Ar LM A
c. Bo nau shaar ema darroiin.
LM Ar A

The verb agrees with the subject in person and number and due to this feature the Actor is often omitted. In these cases, the Landmark is often placed in initial position, as in (11). Further, if the Landmark is animate, it is typically in final position (12), similar to corresponding sentences in English or Swedish.

(11) Bo nau shaar darroiin.
LM A
(12) Mn ktaab-eka-m daa ba Sara.
I book.DEF.1.SG give.PAST to Sara
Ar P A LM
"I gave the book to Sara".

1 When no reference is provided, the example was elicited from a Swedish-speaking native speaker of Sorani-dialect in Malmö, January 2014.

In sum, we have shown here some relevant facts about Swedish and Kurdish word order. Only the proto-roles Ar (Actor), P (Patient), and LM (Landmark) will be of interest.

4 Methods and predictions

We chose to study speakers of Swedish (as spoken in Sweden) and Kurdish (as spoken in the autonomous region of Kurdistan in Iraq) due to the existence of both differences and similarities in the word order of the two languages, as described in Section 3. As shown in the previous section, the differences concern above all the placement of Patient (after the verb in Kurdish, and before in Swedish) while the placement of Landmark is similar.

4.1 Materials

The materials used in the experiments were adapted from those used by Goldin-Meadow et al. (2008), but with some key changes in how the pictures were constructed and presented, as we explain below. The same 36 short video clips were used as stimuli: some were animations using toy figures such as those shown in Figure 1, while others showed live footage of people and animals. In 17 events there was a Patient and in 22 a Landmark. Both Landmark and Patient appeared in eight

Figure 1: The start screen of one of the animated video clips (#27 in Table 2).

events. The actants in each event are specified in Table 2. These are the same classifications as those used by Goldin-Meadow et al. (2008), though here they are consistently described using full sentences and with the three semantic roles: Actor, Patient and Landmark.

Table 2: Video-clips and constituents (actants) analysed in terms of the three semantic roles Actor, Patient and Landmark.

#	Description	Actor	Patient	Landmark
1	*A chicken moves to a captain.*	chicken		captain
2	*A garbage man pushes a garbage can to a cross man.*	garbage man	garbage can	man
3	*A bike moves to a woman.*	bike		woman
4	*A girl topples over.*	girl		
5	*A man moves to a motorcycle.*	man		motorcycle
6	*A boy stirs a spoon in a bowl.*	boy	spoon	bowl
7	*A baby crawls to a chicken.*	baby		chicken
8	*A man bends over.*	man		
9	*A boy moves to a girl.*	boy		girl
10	*A man carries a chicken to a scaffolding.*	man	chicken	scaffolding
11	*A man crawls to a cat.*	man		cat
12	*A captain swings a pail.*	captain	pail	
13	*A gate hits a woman.*	gate	woman	
14	*A girl waves.*	girl		
15	*A man gives a toolbox to a captain.*	man	toolbox	captain
16	*A girl puts down a basket.*	girl	basket	
17	*A girl puts on a hat.*	girl	hat	
18	*A basket moves to a woman.*	basket		woman
19	*A duck moves to a wheelbarrow.*	duck		wheelbarrow
20	*A chicken moves in a circle.*	chicken		
21	*A man picks up a baby.*	man	baby	
22	*A bike with a cart carries a girl to a giraffe.*	bike	girl	giraffe
23	*A dog moves to a van.*	dog		van
24	*A boy tilts a glass.*	boy	glass	
25	*A washer woman pets a dog.*	woman	dog	
26	*A toolbox moves on a school bus.*	toolbox		school bus
27	*A girl gives a flower to a man.*	girl	flower	man
28	*A bike moves to a dog.*	bike		dog
29	*A gate moves to/closes a fence.*	gate		fence
30	*A van drives to a man.*	van		man
31	*A man throws a ball into a basket.*	man	ball	basket
32	*A bicycle bends.*	bicycle		
33	*A man plays a guitar.*	man	guitar	
34	*A dog carries a flower to a doghouse.*	dog	flower	doghouse
35	*A train moves into a fenced area.*	train		fenced area
36	*A woman twists a music box.*	woman	music box	

The first difference in the design compared to the study of Goldin-Meadow et al. (2008) was the way the pictures to be ordered by the participants were constructed. Each image was printed on a white card (10 × 10 cm) using only black colour. Pictures displaying persons were approximately 4–6 cm × 2 cm; pictures of objects 1.5–8 cm × 1.5–8 cm and pictures/symbols of actions (expressed by arrows, as in the original study) were 2–7 cm × 0.5 cm, as shown in Figure 2. The bottom side of each picture was marked with a dark black line, so as to help the participants to interpret the rather schematic pictures, especially the arrows that stood for Acts. When laminated the size of each picture card was approximately 13 × 13 cm. Accordingly, a white cardboard of size 13 × 52 cm with four slots was used to place the pictures on, and participants were instructed to place the pictures below one another, from the slot furthest away from them to the slot closest to them.

Through these modifications we hoped to obtain two main goals. First, since the pictures were not transparent, they had to be placed in a specific order to produce a representation of the event. In other words, the task required a consistent ordering strategy. The second goal was that the participants would be less influenced by the direction of motion displayed in the clips. For example, if an arrow points to the right (as in Figure 3), a participant might choose to place the Actor on the left and the Patient/Landmark on the right side of the arrow, and thus produce an iconic representation of the event. The risk for this would be even higher if there was a Landmark in the direction to which the Actor moved. The vertical layout of the response cardboard was presumed to prevent, or at least minimize, this tendency.

The video clips were arranged in three different random orders, of which each was used equally many times. A laptop with a 15.6 inch screen was used for showing the video clips. Two extra video clips were recorded to use as practice trials, i. e. to help the participants understand the task. Photos of the actors and the objects appearing in the clips were taken and printed, and picture cards were prepared of them in an identical way as the ones used for the actual test.

4.2 Participants

Twenty-seven Swedish participants (13 male) in Lund and Malmö with Swedish as their native language were recruited. They were between 16 and 67 years (mean 31.5). Five had finished elementary school, seventeen had finished secondary school and five had a higher education. The Swedish-speaking participants were not monolingual (as the Kurdish), but it was ensured that

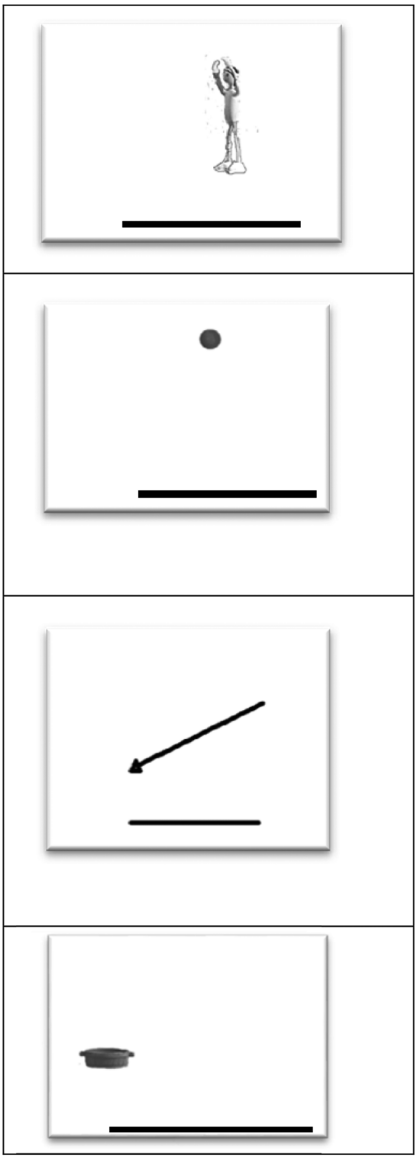

Figure 2: Picture cards used for representing the event: *Man (Actor) throws (Act) a ball (Patient) into a basket (Landmark)* (#31 in Table 2).

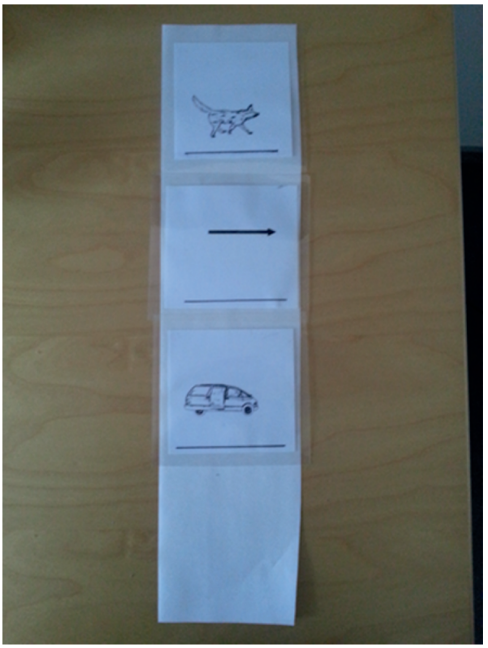

Figure 3: Pictures ordered on the cardboard to represent the event *Dog moves to a van* (#23 in Table 2).

they did not speak a language in which the predominant word order is not SVO. Twenty-six participants (12 male) from Suleymania in the Kurdish region of Iraq were recruited with the help of a Kurdish-speaking assistant. They had their origin in the governorates of Suleymania, Kirkuk and Hawler, and at the time of the study were residing in the cities of Suleymania and Chamchamal. Only monolingual participants[2] in Kurdish were chosen.[3] The ages of the participants were between 16 and 48 (mean 28.6) Eight participants had finished elementary school and thirteen secondary school. Five were illiterate or had very poor literacy skills.

2 The Kurdish and the Swedish groups were not balanced when it comes to level of education, but efforts were taken to select Swedish participants with lower levels of education. It was impossible to find academically educated participants who also were monolingual to perform the task in Kurdistan, thus being monolingual was chosen to be the preferred criterion.

3 By "monolingual," it is meant here that a speaker may know some words or expressions in another language than his/her native tongue, but is not able to communicate in it.

The data from four participants, two in each language group, were discarded. One Kurdish participant was not consistent in arranging the pictures in a top-to-down order, and another used the same order every time as the pictures had when taking them out from the envelope without thinking further. The data of two Swedish participants were discarded because of a failure in video recording. The total number of participants in the dataset, consequently, was 49.

4.3 Procedure

Prior to the testing, forms of informed consent in either Kurdish or Swedish were presented to each participant. Both language-groups were randomly divided into two groups: in the first, they were asked to first describe the video clips verbally and then to arrange the pictures (+*verbalization*); in the second group, they were asked only to use pictures (–*verbalization*). Participants were then interviewed regarding their age, level of education, native language and skills in other languages. Instructions were given both verbally and in writing. The participants were told that they would be shown 36 short video clips. Those in the verbalization condition were instructed to give one-sentence descriptions of "what happened" in the clip. Both groups were told that they would be expected to represent the video clip with a set of pictures, by arranging the pictures on the white cardboard in a top-to-bottom order. Before starting the actual test, participants were given the two training clips and the corresponding pictures. After performing with these, they were asked if they had understood what was expected. If they had questions about the procedure, they were informed without revealing the underlying aim.

Each session was video recorded with a Sony HDR-CX360 Handycam, so that only the stimulus pictures and the hands of the participants were visible. Video clips were shown one by one. If they belonged to the +*verbalization* condition, the participants gave a description of the event verbally and then received the pictures of the constituents in an envelope (in which they were placed in random order). Participants pulled them all out at once and were then allowed to order them on the cardboard in any order preferred. When this was done, the pictures were collected, placed back in the envelope, and the next clip was shown. At the end there was a debriefing session, where participants were asked (a) how they understood the arrow on the pictures, (b) if they had any strategy for using a particular order, and (c) if they could discern what the point of the experiment was.

In Sweden, the experiment was conducted either in a room at the Humanities Laboratory at Lund University (15 cases), at the home of the first author (four cases), in another educational institution (six cases), or at home of a participant (one case). In all situations the participants were alone in the room with the experimenter.

The participant and the experimenter sat at a table beside one another. On the table there was the laptop showing the video-clips by means of a MS PowerPoint presentation, one by one. Beside the laptop was the cardboard for arranging the pictures as well as the video camera on a camera support.

In Kurdistan, the experiments were performed in two houses belonging to the assistant's relatives and in one case the testing was done in the house of a participant. For the first three participants, a table and a chair were obtained. However, since it became clear that participants preferred to do the task on the floor, the rest of the tests were performed so that the participant sat on the floor and had the computer and the cardboard in front of him or her, as shown in Figure 4.[4] The participant then got instructions from the assistant in Kurdish how to perform the task. One assistant showed the video clips on a computer and a second assistant, also a native speaker of Kurdish, handed the pictures after each clip in an envelope, and collected them from the cardboard. Both assistants received exact instructions and they performed consistently. Sometimes the places between the second assistant and the experimenter were changed. Due to the home environment, the participant, experimenter, and two assistants were occasionally not the only people in the room.

Figure 4: Experimental session in Kurdistan: the participant is in the middle, the main assistant to the right, showing the clips on the computer, and the second assistant to the left, handing the picture cards in envelopes.

4 Most people in Kurdistan do all their daily tasks sitting on the floor. They eat, prepare food, drink tea with guests, and watch TV while sitting on the floor.

4.4 Analysis

The information for each participant was made anonymous and saved in a database. The order of pictures used for representing each event was noted by the experimenter by using the codes: Actor (Ar), Act (A), Patient (P) and Landmark (LM). The noted orders were later checked against the video-recordings and if necessary corrected. The software ELAN (tla.mpi.nl/tools/tla-tools/elan/) was used for transcribing the speech of the participants. In the case of Kurdish, transcription was performed by the main Kurdish assistant who was present at the experiment, using a Romanized script developed by the first author.[5] The utterances were then analysed and translated into English by the Kurdish assistant and the first author. The sentence constituents were coded in the same way as for the non-verbal task, using the semantic codes Ar, P, A and LM, as shown in examples (9)-(12). The constituent order information was then saved as an Excel-file, for the purpose of quantitative analysis of preferred constituent orders (a) in all representations, (b) in verbal representations, (c) in non-verbal representations, and (d) whether the order in non-verbal representations correlated with the preceding verbal descriptions.

4.5 Predictions

As described in Section 2.2, Goldin-Meadow et al. (2008) did not find an effect of the dominant word order of the native language of participants on neither the order of placing either the transparencies nor the in the order of the gestural strings. However, our revised design of the pictorial task makes the factor of sequential order more explicit, and thus the task as possibly more analogous to language. Furthermore, as reviewed in Section 2.1, there is an increasing number of studies that show effects of language on non-verbal tasks, especially in contexts that induce "thinking for speaking." Thus, we expected that the two language groups would differ in their relative placement of the pictures representing Patient (P) and Act (A), where the corresponding word order differs, but not in the relative positions A and LM, where it does not. Furthermore, we expected that this effect would be more pronounced in the case where participants had to describe the video-clip prior to representing it with the pictures.

5 There is no standardized Roman script to be used for Sorani Kurdish, which is written with modified Arabic script.

4.6 Results

The complete dataset consisted of 1,764 strings of pictures and 900 spoken utterances. The data were analysed using mixed effects logistic regression. The results that are reported below are the regression coefficients (EST), their standard errors (SE), the *z*-values (*z*) and the associated *p*-values (*p*).

There were 17 items that contained a Patient as well as an Act, yielding a total of 49 × 17 = 833 items for the analysis, 425 Swedish and 408 Kurdish. For the Kurdish group, the predominant picture order was Patient > Act across both conditions, as shown in Table 3, as predicted. In Swedish, the predicted picture order Act > Patient was predominant in the *+verbalization* condition, but not in the *-verbalization* condition.

Table 3: Act-Patient order in the Swedish and Kurdish groups.

	Swedish		Kurdish	
	P > A	**A > P**	**P > A**	**A > P**
–Verbalization	121 (0.59)	83 (0.41)	142 (0.70)	62 (0.30)
+Verbalization	98 (0.44)	123 (0.56)	124 (0.61)	80 (0.39)
	219 (0.52)	206 (0.48)	266 (0.65)	142 (0.35)

A regression model with the predictors, language, and condition was fitted on the proportions given in Table 3. The predicted proportions are shown in Figure 5, supplied with 95% confidence intervals. The interaction of language and condition was not significant, but the main effect of language was (EST = −0.918, SE = 0.453, $z = -2.025$, $p = 0.043$) and the main effect of condition was marginally significant (EST = 0.811, SE = 0.453, $z = 1.793$, $p = 0.073$). In other words, the Kurdish participants used the order A > P significantly less often than the Swedish participants, and the order A > P was more frequent in the *+verbalization* condition than in the *-verbalization* condition in both groups.

In order to follow up on the condition effect, we looked at the linguistic descriptions of the participants in the *+verbalization* condition. Not unexpectedly, the Swedish participants almost always produced an utterance in which the Verb preceded the Object, while the Kurdish participants produced mainly utterances in which the Object preceded the Verb. Nevertheless, in the few cases in which the Swedish participants placed the Object before the Verb in their utterances, they more often placed the Patient picture before the Act picture. The Kurdish participants, on the other hand, placed the Patient picture before the

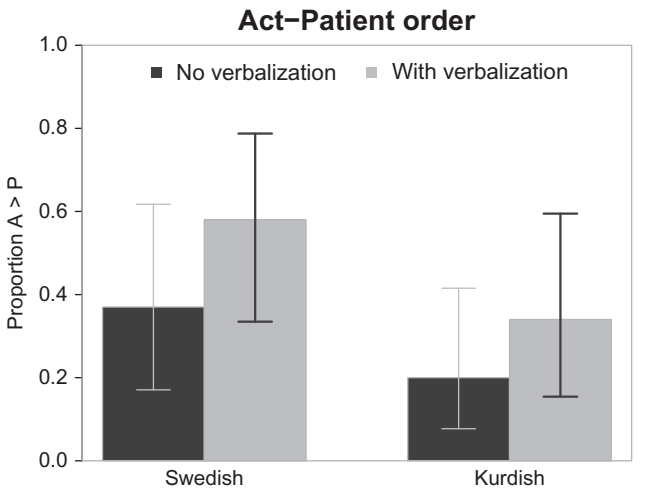

Figure 5: Act > Patient in the Swedish and Kurdish groups; the bars are the predicted proportions from the regression model, the error bars indicate the upper and lower limits of the 95 % confidence intervals.

Act picture more often, irrespective of the order of the constituents in their speech, which indicated less influence of linguistic order on pictorial order than for the Swedish group (Table 4).

Table 4: Order of Act (A) and Patient (P), in relation to Verb (V) and Object (O) in the spoken descriptions *+verbalization* condition.

	Swedish		Kurdish	
	O > V	**V > O**	**O > V**	**V > O**
P > A	9 (0.12)	65 (0.88)	102 (0.95)	5 (0.05)
A > P	3 (0.03)	115 (0.97)	65 (0.96)	3 (0.04)

The data for the relative order between Act and Landmark are presented in Table 5. There were 22 clips with a Landmark picture, yielding $25 \times 22 = 550$ picture sequences produced by the Swedish participants and $24 \times 22 = 528$ picture sequences produced by the Kurdish participants. As expected, the Swedes placed the Landmark predominantly after the Act in the *-verbalization* condition, and even more so in the *+verbalization* condition. However, the Kurdish participants placed the Landmark predominantly before the Act in the *-verbalization* condition.

Table 5: The order of the Landmark (LM) in relation to Act (A) in pictorial representations.

	Swedish		Kurdish	
	LM > A	**A > LM**	**LM > A**	**A > LM**
–Verbalization	73 (0.28)	194 (0.73)	187 (0.71)	77 (0.29)
+Verbalization	17 (0.06)	270 (0.94)	129 (0.49)	134 (0.51)
	90 (0.16)	464 (0.84)	316 (0.60)	211 (0.40)

Interestingly, this changed in the *+verbalization* condition, where they placed the Landmark approximately equally often before as after the Act. A new regression model was also fitted to the proportions given in Table 5. The predicted proportions from the model with their confidence intervals are shown in Figure 6. The interaction of language and condition was not significant, but the main effects of language (EST = 2.580, SE = 1.037, z = 2.487, p = 0.013) and condition (EST = −3.350, SE = 1.007, z = −3.328, p = 0.001) were. In other words, the Swedish participants placed the Landmark after the Act significantly more often than the Kurdish participants, and both groups placed the Landmark after the Act significantly more often in the *+verbalization* condition than in the *-verbalization* condition.

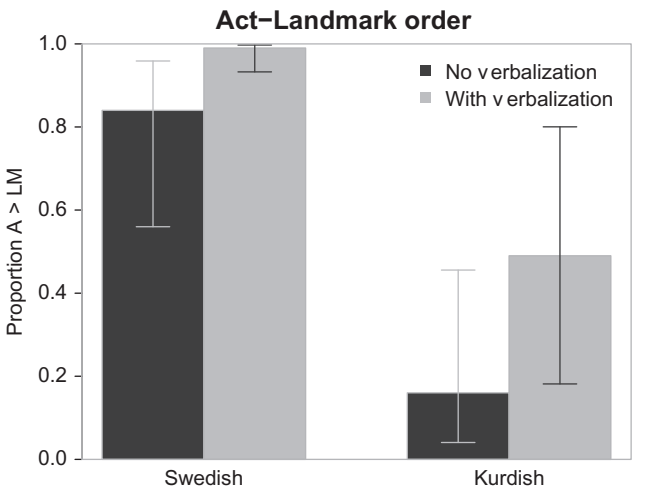

Figure 6: The order of Act (A) and Landmark (LM) in the Swedish and Kurdish groups; the bars are the predicted proportions from the regression model, the error bars indicate the upper and lower limits of the 95 % confidence intervals.

As with the order of Act and Patient, we counted how often the Landmark proceeded and followed the verb in the participants' verbal descriptions. The results showed a very strong tendency for V > LM order in the verbal descriptions for both language groups, and especially for the Swedish participants, but no clear evidence for influence on pictorial order in either group, see Table 6. In sum, while both groups predominantly expressed the Landmark after the verb in their verbal descriptions, the Swedes nearly always placed the Landmark picture after the Act picture, whereas the Kurdish group placed the landmark before the Act picture approximately equally as often as after it.

Table 6: Order of Act (A) and Landmark (LM), in relation to Verb (V) and Landmark in the spoken descriptions in +verbalization condition.

	Swedish		Kurdish	
	LM > V	**V > LM**	**LM > V**	**V > LM**
LM > A	3 (0.30)	7 (0.70)	18 (0.20)	73 (0.80)
A > LM	1 (0.00)	243 (1.00)	11 (0.10)	94 (0.90)

6 Discussion

Our study gave rise to three key findings. First, it should be restated that in the original Golden-Meadow et al. (2008) study, where speakers of four different languages participated, they all behaved similarly with respect to the ways they ordered their non-linguistic representations. In contrast, we found a significant difference in the pictorial orders between the Swedish and Kurdish groups. Importantly, the difference in the order of Act and Patient pictures between the two language groups corresponded to different linguistic word orders, A > P in Swedish, and P > A in Kurdish, supporting the prediction for linguistic influence. This interpretation was supported by the fact that the Swedish participants preferred the A > P order significantly more often after first providing verbal descriptions (56%) than not (41%), see Figure 5. A possible, if still tentative, explanation for the difference between our results and those of Golden-Meadow et al. (2008) is that placing pictures in a row, rather than stacking them on top of one another, is more analogous to linguistic word order – especially with respect for written language.

Second, it should nevertheless be noted that as a whole the Swedish group used both orders almost equally often, indicating that any effect of dominant linguistic word order was counteracted by another factor, biasing for placing the Act after the Patient, such as one of the possible factors discussed in Section 2.2.

Likewise, the fact that the Kurdish group preferred P > A pictorial order could be explained in a similar fashion, and not necessarily as an effect of the dominant word order of their language. How could these two different kinds of factors – the linguistic and the cognitive – be teased apart? We attempted to do so by investigating the relative A-LM order, where both in Kurdish and in Swedish the preferred word order is A > LM. In line with the expected linguistic influence, the Swedish speakers strongly preferred A > LM order in their pictorial representations, in particular in the *+verbalization* condition (see Figure 6). On the other hand, the Kurdish group had LM > A as preferred order, against that in their dominant language, indicating the effect of a cognitive bias. However, they were not completely "immune" to linguistic influence either, as this preference was neutralized in the *+verbalization* condition.

Thus, while our results support the existence of a preference for placing the Act finally, which operates in some contexts and for some kinds of events (as discussed in Section 2.2), the significant difference between the two language groups shows that this preference is only a bias, interacting with other factors, native language word order being one of them. The differences in the results for the *+verbalization* and the *-verbalization* conditions supports this interpretation for the Swedish group and, to some extent, the Kurdish group with respect to the relative order of A and LM.

Third and finally, we are still in need of some explanation, if only a tentative one, for why the relative force for the two determinants (cognitive and linguistic) was different for the two groups. In short, why were the Kurdish participants less influenced by linguistic word order than the Swedes? While we can only answer this question tentatively, there were several indications that pointed in the same direction. We noticed that the Swedish participants took more time in performing the task than the Kurds, indicating more reflective thought processes, and hence possibly more "silent speech." Further, in debriefing they tended to give more elaborate accounts of their strategies, as the following:

> Sometimes I´d think this way: "A man pushes a wagon to another man," and I would use the order man-arrow-wagon-other man. Other times I thought: "A man has a wagon and he takes it to another man," and the order would be: man-wagon-arrow-other man. (Swedish female, -verbalization)

> *I´d put first the thing that was stable and standing in one place, and after that the moving things.* (Swedish female, –verbalization)

In comparison, the Kurdish participants gave few metalinguistic comments in the debriefing, and when pressed to explain how they had solved the task often replied that they did not know, or could not explain. In one case, the following rather idiosyncratic strategy was reported:

> If there was something wrong with the event in the clip, I would put the arrow last, but if the event was normal, I´d place the arrow in the beginning. (Kurdish female, -verbalization)

What the speaker considered "wrong" were cases in which inanimate objects moved on their own, as when a basket moved to a woman without an obvious force causing this.[6]

A possible explanation for these differences could have been cultural. In Kurdistan, everyday life circles around big families and visiting relatives and friends, and less time is spent alone. Thus, the average time used for reading is likely to be less than in Sweden. In addition, five Kurdish participants were overtly illiterate, and some others did not have very good reading and writing skills. As it is well known that literacy effects both language processing, and performance in non-linguistic tasks (Castro-Caldas 1998; Dehaene et al. 2010; Petersson et al. 2001), such differences could help explain why the Kurdish group as a whole appeared less influenced by linguistic word order, and conversely more influenced by the cognitive bias than the Swedish group.

7 Conclusions and future research

In this article, we have addressed a number of cognitive-semiotic issues related to the constituent order of non-verbal event representations. In particular, we took up the classic question concerning the possible influence of language on thought. Representations of motion events using sequences of picture cards were elicited with the help of the same video-clips as those employed by Goldin-Meadow et al. (2008). Unlike in the original study, which did not find any effect of linguistic word order, we used separate picture cards rather than transparencies. These were to be arranged in a top-down order in representing the event seen in the video clips, rather than stacking the transparencies on top of each other. This design change apparently made a difference. There was a significant difference between the pictorial orders used by the Swedish and Kurdish speakers, with the Swedes preferring to place both the Patient and the Landmark pictures after the Act picture in line with the dominant word order in Swedish, and considerably more often than the Kurds. This effect increased after describing the stimuli verbally, supporting a Slobin-style thinking-for-speaking

6 There were some idiosyncratic strategies among the Swedish ones as well. For example, one participant (with some knowledge of Japanese) explained: *I used to read a lot of manga and I was thinking of making a comic of the pictures in the task. I placed them as they would appear in a comic book. I rarely said the actual sentence in my head.* (Swedish male, –verbalization)

hypothesis, or other similar proposals according to which language may mediate thought in some, though not all contexts (Zlatev and Blomberg 2015).

Interestingly, it appeared that the Kurdish group was less influenced than the Swedes by their language, which we believe can be attributed to less familiarity with written language and possibly other cultural factors. However, as the Kurds used the picture order A > LM (in accordance with their word order) twice as often after verbalizing than when not, they were not immune to linguistic influence either.

Thus, our findings support *a joint effect of at least two different factors*: native language word order and a (universal) bias for Actor > Patient > Act order, found in other studies when the events represented are "extensional" and when the events are not "reversible" (Section 2.2): which was the case the present study. The fact that the Swedish participants used this order significantly despite the opposite order in their language supports this interpretation. In sum, a number of different factors appear to impact on the orders of arguments such as Actor, Patient, Landmark, and Act in non-verbal representations. In combination with the factors found in earlier studies, a more complete picture was obtained by adding word order to the list of contributing factors. The major conclusion can be stated shortly: it is neither general cognition nor native language that determine the orders used in non-verbal representations, but rather their interaction. This is admittedly a more complex explanation than either-or, but the only one that is fully consistent with the findings. Future studies will need to elaborate this empirically and theoretically.

We conclude by outlining two other semiotic issues that such studies should take into consideration. One has to do with the perceived "realism" of the medium in which events are presented. In our study, we noticed that the Kurdish participants described the animated characters in the video-clips with terms that were appropriate for the real-world referents, while the Swedes referred to intermediate representations. For example, while the Kurds spoke of "a man", "a dog" and "a girl", the Swedish speakers denoted the same animated characters as "a doll" or "a toy". This skipping of the intermediary semiotic level may have to do with different generalizations from pictures to real objects between cultures with different degrees of exposures to pictures (Walker et al. 2013). Alternatively, the difference in referential terms could also depend on the fact that Kurds have had a much shorter history of providing industrially produced objects for playing to their children.

A second issue concerns the pictorial representations of Acts. The debriefing sessions in our study showed that there was not much variation in how the pictures of the Actor, Patient or Landmark were interpreted. At the same time, the responses varied considerably as to how the schematic Act-pictures were understood. The different arrows were interpreted as motions, actions, directions,

and labelled with verbs like "take", prepositions like "to", and "towards" or directional adverb like "there." In other words, while all pictures were predominantly iconic representations, this iconicity was *diagrammatic* in the case of Acts and *imagistic* for all other pictures (Peirce 1931–1958). This difference could possibly lead to a confound in ordering tasks, since the placement of verbs and other form classes like prepositions and adverbs differs in both Swedish and Kurdish. A possible solution in future research could be to utilize pictures with a higher grade of imagistic detail for the Act representations. For instance, these could be images of a hand along with an arrow to represent "twist," or a picture of legs in movement to express "walk," relying on a combination of iconicity and indexicality.

These additional questions, as well as the main topic of this field of research (as well as our empirical findings) – the interaction between linguistic and other cognitive factors over sequential non-verbal event representations – clearly indicate the relevance of the field for cognitive semiotics.

Acknowledgements: We would like to thank Prof. Goldin-Meadow for allowing us to use the video-clips and pictures used in the original study, and to Jennifer E. Arnold and an anonymous reviewer for helpful comments.

References

Casasanto, D., L. Boroditsky, W. Phillips, J. Greene, S. Goswami, S. Bocanegra-Thiel, et al. 2004. How deep are effects of language on thought? Time estimation in speakers of English, Indonesian, Greek, and Spanish. In K. Forbus, D. Gentner & T. Regier (eds.), *Proceedings of the 26h Annual Conference Cognitive Science Society*, 575–580. Hillsdale: Lawrence Erlbaum Associates.

Casasanto, D. 2008. Who's afraid of the big bad Whorf? Cross-linguistic differences in temporal language and thought. *Language Learning* 58. 63–79.

Castro-Caldas, A., K. M. Peterson, A. Reis, S. Stone-Elander & M. Ingvar. 1998. The illiterate brain. Learning to read and write during childhood influences the functional organization of the adult brain. *Brain*, 121(6). 1053–1063.

Christensen, P. & T. Kristian. 2013. Gestural representation of event structure in dyadic interaction. Tilburg Gesture Research Meeting 2013. http://tiger.uvt.nl/list-of-acceptedpapers.html.

Christensen, P., F. Riccardo & T. Kristian. 2016. Environmental constraints shaping constituent order in emerging communication systems: Structural iconicity, interactive alignment and conventionalization. *Cognition* 146. 67–80.

Dehaene, S., F. Pegado, L. W. Braga, P. Ventura, G. N. Filho, A. Jobert, G. Dehaene-Lambertz, R. Kolinsky, J. Morais & L. Cohen. 2010. How learning to read changes the cortical networks for vision and language. *Science* 1359. 330. DOI: 10.1126/science.119414.

Dixon, R. M. W. 2012. *Basic linguistic theory: Vol. 3. Further grammatical topics*. Oxford: Oxford University Press.
Dowty, D. 1991. Thematic proto-roles and argument selection. *Language* 67. 547–619.
Dryer, M. S. 2013. Order of subject, object and verb. In M. S. Dryer & M. Haspelmath (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. http://wals.info/chapter/81. Accessed 9 September 2013.
Fodor, J. A. 1983. *Modularity of mind*. Cambridge: MIT Press.
Gell-Mann, M. & M. Ruhlen. 2011. The origin and evolution of word order. *PNAS* 108(42). 17290–17295.
Gibson, E., L. Bergena & S. T. Piantadosic. 2013. Rational integration of noisy evidence and prior semantic expectations in sentence interpretation. *PNAS* 110. 20.
Givón, T. 1979. *On understanding grammar*. New York: Academic Press.
Givón, T. 1990. *Syntax: a functional-typological introduction, Volume 1 and 2*. Amsterdam: John Benjamins.
Goldin-Meadow, S., C. Mylander, W. C. So & A. Özyürek. 2008. The natural order of events: How speakers of different languages represent events nonverbally. *PNAS* 105. 9163–9168.
Haig, G. & Y. Matras. 2002. *Kurdish linguistics: A brief overview*. Berlin: Sprachtypol. Univ. Forsch. (STUF).
Hall, M. L., R. I. Mayberry & V. S. Ferreira. 2013. Cognitive constraints on constituent order: Evidence from elicited pantomime. *Cognition* 129. 1–17.
Humboldt, W. 1988. On language: The diversity of human language-structure and its Influence on the mental development of mankind. Cambridge: Cambridge University Press.
Josefsson, G. 2004. Non-target structures and non-target uses in child and adult Swedish. In G. Josefsson, C. Platzack, G. Håkansson (eds.), *The acquisition of Swedish grammar*, 155–193. Amsterdam: John Benjamin.
Kita, S. & A. Özyürek. 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. 16–32.
Langus, A. & M. Nespor. 2010. Cognitive systems struggling for word order. *Cognitive Psychology* 60. 291–318.
Levinson, S. C. 2003. *Space in language and cognition. Explorations in cognitive diversity*. Cambridge: Cambridge University Press.
Majid, A., M. Bowerman, S. Kita, D. M. B. Haun & S. C. Levinson. 2004. Can language restructure cognition? The case for space. *Trends in Cognitive Sciences* 8(3). 108–114.
Meir, I., A. Lifshitz, D. Ilkbasaran & C. Padden. 2010. The interaction of animacy and word order in human languages: A study of strategies in a novel communication task. *The Evolution of Language: Proceedings of the 8th International Conference*, 455–456.
McNeill, D. 2000. Analogic/analytic representations and cross linguistic differences in thinking for speaking. *Cognitive Linguistics* 11(1–2). 43–60.
Núñeza, R. E. & E. Sweetser. 2006. With the future behind them: Convergent evidence from Aymara language and gesture in the crosslinguistic comparison of spatial construals of time. *Cognitive Science* 30. 401–445.
Peirce, C. S. 1931–1958. *Collected papers of Charles Sanders Peirce*. In C. Hartshorne and P. Weiss, eds. volumes 1–6 and A. W. Burks, ed. volumes 7–8. Cambridge: Harvard University Press.

Petersson, K. M., A. Reis & M. Ingvar. 2001. Cognitive processing in illiterate and literate subjects: A review of some recent behavioural and functional neuroimaging data. *Scandinavian Journal of Psychology* 42. 251–267.
Pinker, S. 1994. *The language instinct*. London: The Penguin Press.
Saeed, J. I. 2009. *Semantics*. Chichester: Wiley-Blackwell.
Sandler, W. 2010. The uniformity and diversity of language: Evidence from sign language. *Lingua* 120.
Schouwstra, M. 2012. *Semantic structures, communicative strategies and the emergence of language*. LOT publications 312. PhD thesis, Utrecht Institute for Linguistics. OTS.
Schouwstra, M. & H. de Swart. 2014. The semantic origins of word order. *Cognition* 131. 431–436.
Slobin, D. I. 1996. From "thought and language" to "thinking for speaking". In J. J. Gumperz & S. C. Levinson (eds.), *Rethinking linguistic relativity*, 70–96. Cambridge: Cambridge University Press.
Talmy, L. 2000. Toward a cognitive semantics: Concept structuring systems. Cambridge, MA: MIT Press.
Thackston, W. M. 2006. *Sorani Kurdish, A Reference Grammar with Selected Readings*. http://www.fas.harvard.edu/~iranian/Sorani/sorani_1_grammar.pdf (accessed 19 March 2013).
Walker, C. M., L. B. Walker & P. A. Ganea. 2013. The role of symbol-based experience in early learning and transfer from pictures: Evidence from Tanzania. *Developmental Psychology* 49(7). 1315–1324.
Whorf, B. L. 1956. Science and linguistics. In J. B. Carroll (ed.), *Language, Thought and reality: Selected writings of Benjamin Lee Whorf,* 207–219. Cambridge: MIT Press.
Zlatev, J. & J. Blomberg. 2015. Language may indeed influence thought. *Frontiers in Psychology, Cognitive Science* 6. 1631. DOI: 10.3389/fpsyg.2015.01631

Bionotes

Anu Vastenius

Anu Vastenius is a doctoral student at the Division for Cognitive Semiotics, Centre for Languages and Literature, Lund University. Her research interests are experimental semiotics, linguistic relativity, gesture studies, and language evolution.

Joost van de Weijer

Joost van de Weijer obtained his Ph.D. from the Max Planck Institute for Psycholinguistics in 1998. He is currently employed at the Department of Linguistics and Phonetics and at the Laboratory of the Humanities, both at Lund University. His main research interests are experimental linguistics and quantitative data analysis.

Jordan Zlatev

Jordan Zlatev is Professor of General Linguistics, working at the Division for Cognitive Semiotics, Centre for Languages and Literature, Lund University. His research focuses on language as a predominantly conventional semiotic system for communication and thought in relation to other semiotic systems and consciousness. He is the author of Situated Embodiment: Studies in the Emergence of Spatial Meaning (1997), co-editor of The Shared Mind (2008) and Moving Ourselves, Moving Others (2012), and has published over 60 articles in peer-reviewed journals and books.