Semantic structure in improvised communication

Marieke Schouwstra (Marieke.Schouwstra@phil.uu.nl)
UiL OTS, Utrecht University
Trans 10, 3512 JK, Utrecht, the Netherlands

Anouschka van Leeuwen (A.vanLeeuwen@uu.nl)
Nicky Marien (nickymarien@hotmail.com)
Marianne Smit (mariannesmit@gmail.com)
Henriëtte de Swart (H.deSwart@uu.nl)
Utrecht University
Utrecht, the Netherlands

Abstract

Recent studies have shown that people, when asked to communicate about simple events in an improvised manner, i.e., using only gesture and no speech, consistently use a sequencing that corresponds with SOV order. We present experimental data showing that SOV is not the only order that emerges, and that *intensional events* give rise to a different ordering: SVO. We conclude that the semantic properties of the events that are communicated about play a role in the sequencing of utterances in emerging language systems. Further, we hypothesise that in simple language systems the sequencing of elements has a communicative function. We present a second study, which shows that different ordering of the constituents in improvised communication sequences results in different interpretations.

Keywords: gesture; language genesis; semantics; word order

Introduction

In Goldin-Meadow, So, Özyürek, and Mylander (2008), it is investigated how people sequence information when they are asked to communicate about simple events using gesture and no speech. Speakers of four languages with different dominant word orders (English (SVO), Chinese (SVO/SOV), Spanish (SVO) and Turkish (SOV)) were asked to describe simple events depicted on vignettes using only gesture and no speech (none of the participants were familiar with any conventional sign language). Each vignette depicted a *motion event*: a simple event containing an actor, a patient and an act that is generally described with a transitive sentence containing a subject, a direct object and a verb, and in which the act involves motion (e.g. 'girl covers box' or 'captain swings pail').

It was found that the participants, although they were speakers of languages with different dominant word orders, used a consistent order for their gesturing: Actor-Patient-Act. Goldin Meadow et. al. point out that this order, Actor-Patient-Act corresponds with the sentence order SOV. They conclude that Actor-Patient-Act (henceforth, SOV)[1] may reflect a natural sequencing for representing events.

In Langus and Nespor (2010), it was shown that also in gesture *comprehension*, SOV order is preferred: it led to the shortest reaction times in a gesture comprehension task carried out by Turkish (SOV) and Italian (SVO) participants. Thus, the word order of one's native language is bypassed in both the production and interpretation of improvised communication, and SOV ordering is preferred instead.

[1] Following Langus and Nespor (2010), we will use SOV to refer to Actor-Patient-Act.

Motion events vs. intensional events

Goldin Meadow et al speculate that SOV 'may reflect a natural sequencing for representing events.' As an explanation for this particular order they quote results from related research, showing that the Subject and Object might be situated before the Verb, because entities are cognitively more basic and less relational than actions Gentner and Boroditsky (2001). Moreover, objects and actions are cognitively tied Goldin-Meadow (2003), which would link Object to Verb, and result in the SOV order for gesturing (Goldin-Meadow et al., 2008, p. 9166).

This explanation seems intuitive, especially for the particular kind of events that were used in the experiment described in Goldin-Meadow et al. (2008). We found, however, that there is a category of events, *intensional events*, for which the SOV order seems less intuitive. Intensional events differ from motion events semantically and we will hypothesise that different semantic properties lead to different gesture orderings in the improvised communication task. But first, let us focus on the differences between motion events and intensional events.

The events used in (Goldin-Meadow et al., 2008) are all events in which someone does something to someone or something else. In these situations, the ontological status of subject and direct object are similar. E.g., in the example of a girl covering a box, we can summarise the situation as follows: (1) there is a girl, (2) there is a box, and (3) the girl covers the box. In other words, in order for the sentence to be true, both the subject and the direct object need to exist, and they need to relate to each other in the right way. Let us compare this situation with the following example: 'a princess wants an apple'. In this example, the ontological status of the subject and the direct object are not equal: in order for the sentence to be true, we need the princess to exist, but the 'ontological demands' on the apple are different: a princess can want an apple without the actual apple being around, or she can want an apple but not one in particular. It is even possible for the princess to want something that does not exist at all.

The crucial difference between the two events thus resides in a difference between the *verbs* that describe the actions going on in them: both 'cover' and 'want' are transitive verbs (they occur with a subject and a direct object), but 'cover' is an extensional verb, whereas 'want' is an *intensional* verb. Other examples of intensional verbs are 'seek', 'admire', and arguably also 'see' and 'draw'. In the literature, interesting properties of intensional verbs have been described and inventories of intensional verbs have been made (Forbes, 2010; Moltmann, 2008). The terms 'intensional' vs 'extensional' are used because of the role that extensions (the object a term refers to) and intensions (the meaning of a term) play in the interpretation of these verbs. In order to interpret an extensional verb, the extension of its complement (the direct object) is important, whereas for the interpretation of intensional verbs, the extension of its complement is less important than its meaning. For a precise characterisation of intensional verbs, see the Appendix.

Let us look again at the event 'princess wants apple'. This event is typically described with an intensional verb, and can thus be called an intensional event. As pointed out, there is something special about the direct object in such events: the apple in the 'want-event' is not necessarily a concrete object. If we would make a step-by-step analysis of the event, it would look like this: (1) there is a princess, (2) there is something the princess wants, and (3) that is an apple. This analysis reflects the fact that there is something special about the direct object: it is in some sense dependent on the Subject and the Verb.

Intensional events and gesturing order

We have described semantic differences between two kinds of events. How will these two kinds of events behave in the improvised communication setting? An interesting link between semantic properties and word order is provided in Jackendoff (2002). Jackendoff suggests that in simple language systems without full syntax, semantic principles might play an organising role in short utterances. We will assume that the gesture strings produced in the improvised communication task are such a 'language system without full syntax', and thus hypothesise that *semantic properties of events are important in the improvised communication setting and will influence the order of the gesturing.*

To test this hypothesis, we set up an improvised communication experiment where people are asked to convey the meanings of both motion events and intensional events. We predict that for motion events, which are all typically described by extensional verbs, participants will use SOV order, similarly to what was shown in (Goldin-Meadow et al., 2008). In intensional events, however, there is something exceptional about the ontological status of the direct object. We have seen above that the direct objects in such events are less concrete, and we predict that they will be placed in the end of the gesturing sequence, resulting in a gesturing order of SVO for intensional events.

Experiment: Gesturing motion events and intensional events

Method

16 participants (5 male, 11 female) were recruited from Utrecht University and the Utrecht School of the Arts. All were native speakers of Dutch (which is an SVO language), and none of the participants had any knowledge of a conventional sign language.

The set of items consisted of 20 pictures of motion events (e.g. 'Pirate throws guitar', 'Princess carries vase'), and 20 pictures of intensional events (e.g. 'Cook thinks of sock', 'Leprechaun sees tall building'). Each motion event had a corresponding intensional event, with the same actor and patient, but a different action. All actors (subjects) in the pictures had particular external characteristics (e.g. a princess with a crown, a pirate with a hat), in order to encourage participants to gesture all elements in the picture. All patients (direct objects) were inanimate objects.[2] All pictures were pre-tested for clarity. Each picture was shown either in its original version or as a mirror image, to control for the left-to-right order of the elements in the pictures.

For the experiment, two versions were created, each consisting of 10 pictures of motion events and 10 pictures of intensional events. The items were presented in random order. Participants were shown pictures of events on a computer screen. They were asked to convey the meaning of each picture to the experimenter (who could not see the computer screen), by using only gestures and no speech. Each picture remained visible on the screen while the participant was gesturing. Participants were told to keep gesturing until they thought they had conveyed the meaning of the picture; no information was given about the amount of gestures to be used.

Before the actual experiment started, participants were shown four practice items. During the practice stage of the experiment, the experimenter gave feedback about whether or not she understood which meaning was conveyed. No spoken feedback was given during the experiment.

After the gesturing part of the experiment, participants were shown the pictures again, and were asked to describe each event using a Dutch sentence.

Data analysis and results

The video recordings were coded for gesturing order by two independent coders (80,6% agreement). All gesturing sequences for which there was no consensus were filtered out (62 of 320 recordings). Occasionally, participants produced gesture strings describing an action that did not match with the intended action on the picture; these were removed as well (10 recordings).

The overall results are shown in figure 1. The chart shows the percentages of SOV strings, SVO strings and

[2]This was done to exclude any effects of animacy described in (Meir, Lifshitz, Ilkbasaran, & Padden, 2010).

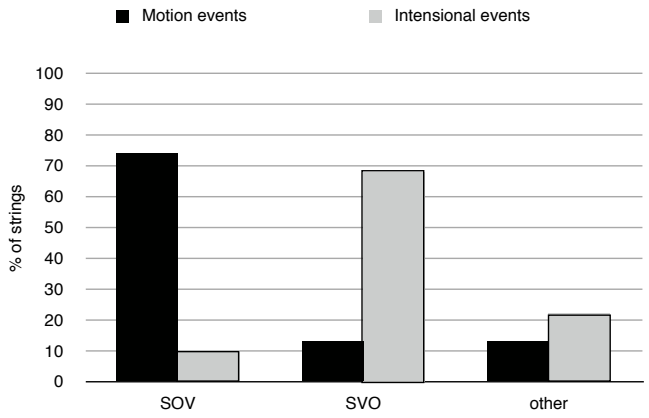

Figure 1: Results of the production experiment

other orders. The category of 'other' orders (which is only a minority of the totality of strings) consisted of strings like OSV, VSO, or strings with either less or more than three gestures.

The data was analysed using a repeated measures ANOVA. The within subject factors were Picture-type (intensional or motion) and Order (SOV or SVO); the between subjects factor was Version (version 1 or 2). Among 132 gesture strings of motion events, the proportion of SOV order was high ($M = .735$, $SE = .065$), whereas the proportion of SVO order was low ($M = .132$, $SE = .069$). Among 122 gesture strings of intensional events, the proportion of SVO order was high ($M = .649$, $SE = .056$), whereas the proportion of SOV order was low ($M = .103$, $SE = .037$).

We found a significant interaction between Picture type and Order: $F(1,14) = 100.753$, $p = .000$. No significant interaction was found between the main interaction and the effect of version ($F(1,14) = .008$, $p = .932$).

Discussion

The results were as expected: for motion events, the results obtained in Goldin-Meadow et al. (2008) were replicated, but for intensional events, different gesture sequencing was used: the direct object was placed after the verb, resulting in an SVO order. This confirms our hypothesis that semantic properties influence ordering in improvised communication.

Pilot study: interpreting SOV and SVO strings

It has become clear from the experiment described above that the semantic differences between motion events and intensional events have an influence on gesture production in the improvised communication task.

We hypothesise that these different orderings serve a communicative purpose. Given that communication involves both the production of strings and their interpretation, and that we have already seen the effects of semantic differences on the production side, we need to find out whether different gesturing orders have an influence on the interpretation of these gestures. We set up an experiment with a series of ambiguous gesturing sequences in two orders: SOV and SVO. If our hypothesis that the usage of different ordering for different semantic structure serves a communicative purpose is true, we expect SOV strings to be interpreted as motion events and SVO strings as intensional events (van Leeuwen, 2010).

Method

We created short movie clips showing an actor gesturing simple events. The verbs were gestured in such a way that the events acted out could be interpreted either as a motion event or as an intensional event.

Figure 2: An ambiguous action: 'climb' or 'build'.

An example of the way in which these ambiguous items were gestured is given in figure 2, in which the ambiguous action 'climb/build' is shown. Thus, the gesture in the figure can be interpreted as a climbing action, as well as a building action. We created two videos of each ambiguous verb by adding a gestured subject and object to the transitive event. There were two different orders: [Subject-AmbiguousVerb-Object] and [Subject-Object-AmbiguousVerb]. Thus, each video consists of *exactly the same* video material, but the elements are put in two different orders.

The following ambiguous events were used; for every item in the list, the option marked with *m* creates a motion event, and the option marked with *i* creates an intensional event.

- Pirate drops$_m$/searches$_i$ ball.
- Princess breaks$_m$/sculpts$_i$ vase.
- Leprechaun cuts$_m$/draws$_i$ pizza.
- Witch eats$_m$/wants$_i$ banana.
- Witch paints$_m$/paints$_i$ table.[3]
- Girl sleeps on$_m$/dreams of$_i$ book.
- Girl kisses$_m$/thinks of$_i$ doll.
- Princess talks to$_m$/talks about$_i$ snowman.
- Pirate throws$_m$/hears$_i$ guitar.
- Cook stirs$_m$/smells$_i$ soup.

[3] In the first interpretation a witch painting an existing table is meant; in the second a witch painting a table on a canvas.

- Leprechaun hits$_m$/feels$_i$ book.
- Witch climbs$_m$/builds$_i$ house.

Of the 12 video pairs, two versions were created, each consisting of 6 videos in SVO order, 6 videos in SOV order. Four fillers, items with unambiguous actions, were added to each version. The videos were shown to participants in a two alternative forced choice task; pictures of the corresponding intensional and extensional (motion) events were shown as the two answer possibilities.

Forty one native speakers of Dutch were recruited from the Utrecht University library (they did not receive a monetary compensation). They were shown videos on a laptop screen and were asked to choose, after each video, the picture that fitted best with the event acted out on the video. First two practice items with unambiguous verbs were shown, followed by the ambiguous items and fillers (ambiguous items and fillers were presented in random order). The two answer possibilities were shown before each video, and again afterwards (the order of the two answer possibilities was also randomly determined).

Data analysis and results

Upon re-analysis of the video clips we decided to exclude two videos from the results: 'Pirate drops$_m$/searches$_i$ ball' and 'Girl kisses$_m$/thinks of$_i$ doll'. These two videos differ from the others in the sense that the ambiguous actions they depict consist of two sub-gestures,[4] whereas for all other ambiguous actions, only one gesture is used.

The data was analysed using a repeated measures ANOVA. The within subjects factor was Order (SOV or SVO), and the within subjects factor was Version (version 1 or 2). The results are shown in figure 3

For SOV gesturing sequences, a motion answer was chosen more often ($M = .751$, $SE = .020$) than for SVO gesturing sequences ($M = .488$, $SE = .032$). We found a significant main effect of Order: $F(1,39) = 42.709$, $p = .000$. No significant interaction was found of version ($F(1,39) = .007$, $p = .934$).

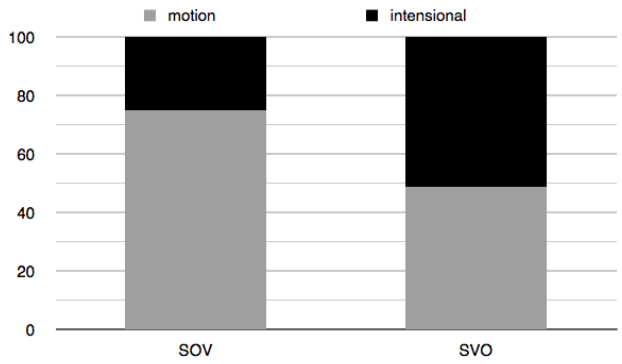

Figure 3: Results of the interpretation study

[4] A drop-gesture followed by a search gesture for the former, and a think of gesture followed by a kiss gesture for the latter.

Discussion

The two different gesturing orderings led people to interpret the gesturing differently: SOV gesturing strings were more likely to be interpreted as motion events than SVO strings were, and SVO gesturing strings were more likely to be interpreted as intensional event than SOV strings were. So we are safe to conclude that the order of gesturing has an influence on the interpretation of the gesture strings. This supports our hypothesis that choosing a different ordering in order to convey different kinds of meanings serves a communicative purpose.

General discussion

In previous publications, similarities were pointed out between the improvised communication task and newly emerged sign languages like Al Sayyid Bedouin Sign Language and Nicaraguan Sign Language.[5] And from these similarities it was hypothesised that results from the improvised communication experiment can tell us something about the emergence of language in general: it was argued that SOV word order *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, p. 9167). This in turn connects well to the claim made in (Newmeyer, 2000), that the earliest human language had rigid SOV word order.

The production experiment described in this paper shows that it is not SOV word order as such that is important in the improvised communication task, but rather the meanings that are to be conveyed. It was shown that for gesture production, different event types were gestured in different orders.

The interpretation study described in this paper shows that the order of gesturing has an influence on the interpretation of the gesture strings. This supports the hypothesis that the distinction found in gesture production between SVO and SOV ordering for intensional events and motion events respectively, has a communicative function.

Effects of meaning on structure in simple language systems have been found in other linguistic phenomena where improvisation is required. In the process of acquiring a second language outside the classroom, adult learners go through a stage that has been characterized as being (1) determined by a small number of organisational principles, (2) largely independent of the source or target language of the learner and (3) simple but relatively successful for communication. This stage is called the Basic Variety (Klein & Perdue, 1997). Some examples of organisational principles of the Basic Variety are FocusLast ('put the information that is in focus, new information, at the end of the sentence') and AgentFirst ('the NP referent with the highest control comes first'). Similar organisational principles were described for, e.g., pidgin languages (Jackendoff, 2002).

The fact that the organisational principles described above are found consistently in linguistic phenomena like the Ba-

[5] See (Goldin-Meadow et al., 2008, p. 9167).

sic Variety and pidgins, and seem to be independent of the native language of their users, makes these phenomena interesting for the language evolution debate: they might tell us something about the structure of evolutionarily early language (Jackendoff, 2002; Schouwstra, 2010).

The circumstances under which people create systems like the Basic Variety and pidgin are very particular and also often undesirable. Moreover, the circumstances under which these restricted systems emerge are not controllable, and the data is therefore not very clean. Finally, the focus of data collection for these systems was mainly on production, and not on comprehension. Therefore, it would be valuable to able to collect this kind of data in a controlled environment, like in a laboratory. The improvised communication task provides us with exactly this: if we see the task as a setting where restricted linguistic systems are produced, we can obtain data in the ideal way described above.

The improvised communication task can be seen as an environment where restricted linguistic systems are produced, because the setting shares many properties with the settings of pidgin and unsupervised second language learning. In the experiment, like in those situations, subjects cannot use their native language to express themselves and are forced to improvise, using whatever they have in their restricted inventory. The improvised communication task thus offers us a way to collect data about the earliest stages of language emergence in a controlled manner.

Moreover, the view of the improvised communication task as a restricted linguistic system allows us to compare the influence of semantic structures in existing studies into, e.g., the Basic Variety and pidgin languages to that in lab situations.

Conclusion

In this paper, we have shown that when people are forced to communicate in an improvised manner, this does not necessarily lead to SOV ordering of utterances. Rather, the semantic properties of verbs are decisive in the ordering of gestures in the improvised communication task. This supports the view that in language systems without full syntax, semantic properties play an organising role, a process that is also seen in other linguistic situations where improvisation is required (e.g. early stages in spontaneous second language acquisition and pidgin).

We have also shown that different ordering of the constituents in improvised communication sequences results in different interpretations. This supports the view that choosing different utterance structures in order to express different kinds of meanings has a communicative function.

Appendix: intensional transitive verbs

In order to get a more precise characterisation of the differences between extensional and intensional transitive verbs, we will give a brief overview of three 'marks of intensionality that were described in (Forbes, 2010): substitution-resistance, the availability of unspecific readings, and existence-neutrality.

Substitution-resistance

In sentences with extensional verbs, it is possible to substitute the direct object with one that refers to the same object, without changing the truth value of the sentence. This is illustrated in the following examples:

(1) John lives next to Mark Twain.

(2) John lives next to Samuel Clemens.

Because Samuel Clemens is Mark Twain, sentence 1 is true in exactly the same situations as sentence 2. If we substitute the extensional verb live next to for an intensional verb, admire, this is no longer possible.

(3) John admires Mark Twain.

(4) John admires Samuel Clemens.

It might be the case that sentence 3 is true, but that John does not realise that his grumpy neighbour Samuel Clemens is Mark Twain. In that case, 'John admires Mark Twain is true, while 'John admires Samuel Clemens is false.

The availability of unspecific readings

In sentences with intensional verbs, it is possible that the direct object remains unspecific. An example of this is the following sentence:

(5) Mary seeks a man.

For this sentence, an interpretation is possible where Mary seeks a man, but not one man in particular. Contrast this with the verb kiss. We cannot say

(6) Mary kissed a man, but not one in particular.

Existence-neutrality

In sentences with intensional verbs, it is possible for the direct object not to exist at all. By contrast, for extensional verbs, the direct object needs to exist in order for the sentence to make sense. It is possible to *seek* a unicorn, but not to *stumble across* one.

Intensional items used in the experiment

Intensional verbs will always show at least one of the three properties described above. Some verbs manifest all three kinds of behaviour, but there are many verbs that meet only one of the criteria. For certain verbs, it is not always clear whether a particular criterion is met.

The verbs that were used in the production experiment are the following (behind each subset, a name for the subclass of verbs is given):

- search ('classical' intensional verb)
- dream of, think of (psych verbs)
- hear, see (perception verbs)
- build, draw, knit, sculpt (creation verbs)

The class of intensional verbs is thus a rather diverse group and there is no general agreement on either the names of the subcategories, the verbs that should be included in them, or even whether all categories listed above are truly intensional.

One might, for example, question the intensional properties of perception verbs: if John *sees* a house, doesn't he just see an existing external object? But the fact that a sentence like 'When John listened to a cello, he heard a violin' is possible, shows that there is, after all, something special about complements of perception verbs. In support of this view, see the following characterisation of perception verbs, as presented in Moltmann (2008):

> The complements of perception verbs [...] do not describe the external object that may be perceived, but rather the way the perceived object appears.

In other words, for perception verbs, like for other intensional verbs, it is not the external object that is important for their interpretation, but the intension. This shows that perception verbs have at least an intensional flavor.

For our experiments, we have taken the widest definition of 'intensional' and included all verbs that have an intensional flavor.

Acknowledgements

Many thanks to Simon Kirby, Mónica Tamariz and Marijn Struiksma for their insightful suggestions.

References

Forbes, G. (2010). Intensional transitive verbs. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2010 ed.).

Gentner, D., & Boroditsky, L. (2001). Individuation, relativity and early word learning. In M. Bowerman & S. Levinson (Eds.), *Language acquisition and conceptual development* (pp. 215–256). New York: Cambridge University Press.

Goldin-Meadow, S. (2003). *Resilience of language*. New York: Psychology Press.

Goldin-Meadow, S., So, W. C., Özyürek, A., & Mylander, C. (2008). The natural order of events: How speakers of different languages represent events nonverbally. *PNAS*, *105*(27), 9163–9168.

Jackendoff, R. (2002). *Foundations of language: Brain, meaning, grammar, evolution*. Oxford University Press.

Klein, W., & Perdue, C. (1997). The Basic Variety (or: couldn't natural languages be much simpler?). *Second Language Research*, *13*, 301–347.

Langus, A., & Nespor, M. (2010). cognitive systems struggling for word order. *Cognitive Psychology*, *60*(4).

Leeuwen, A. van. (2010). *Gesturing orders and their interpretation*. Unpublished master's thesis, Utrecht University.

Meir, I., Lifshitz, A., Ilkbasaran, D., & Padden, C. (2010). The interaction of animacy and word order in human languages. In A. Smith, M. Schouwstra, B. de Boer, & K. Smith (Eds.), *The evolution of language* (pp. 455–456). Singapore: World Scientific.

Moltmann, F. (2008). Intensional verbs and their intentional objects. *Natural Language Semantics*, *16*(3), 239–270.

Newmeyer, F. J. (2000). On the reconstruction of 'proto world' word order. In C. Knight, J. R. Hurford, & M. Studdert-Kennedy (Eds.), *The evolutionary emergence of language.* Cambridge: Cambridge University Press.

Schouwstra, M. (2010). Language evolution: the view from second language learners. In A. Smith, M. Schouwstra, B. de Boer, & K. Smith (Eds.), *The evolution of language.* Singapore: World Scientific.