Brief article

The semantic origins of word order

Marieke Schouwstra [a,b,*], Henriëtte de Swart [a]

[a] Utrecht Institute of Linguistics OTS, Utrecht University, Trans 10, 3512 JK Utrecht, The Netherlands
[b] School of Psychology, Philosophy and Language Sciences, University of Edinburgh, Edinburgh EH8 9AD, United Kingdom

ARTICLE INFO

ABSTRACT

Where do the different sentence orders in the languages of the world come from? Recently, it has been suggested that there is a basic sentence order, SOV (Subject–Object–Verb), which was the starting point for other sentence orders. Backup for this claim was found in newly emerging languages, as well as in experiments where people are asked to convey simple meanings in improvised gesture production. In both cases, researchers found that the predominant word order is SOV. Recent literature has shown that the pragmatic rule 'Agent first' drives the preference for S initial word order, but this rule does not decide between SOV and SVO. This paper presents experimental evidence for grounding the word order that emerges in gesture production in semantic properties of the message to be conveyed. We focus on the role of the verb, and argue that the preference for SOV word order reported in earlier experiments is due to the use of extensional verbs (e.g. *throw*). With intensional verbs like *think*, the object is dependent on the agent's thought, and our experiment confirms that such verbs lead to a preference for SVO instead. We conclude that the meaning of the verb plays a crucial role in the sequencing of utterances in emerging language systems. This finding is relevant for the debate on language evolution, because it suggests that semantics underlies the early formation of syntactic rules.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

Many languages have a basic ordering of the subject (S), object (O), and the verb (V), and among the languages of the world, all six possible basic word orders exist. Of these six orders, SVO and SOV make up a large majority of the totality of languages (Dryer, 2011). Recent work on the origins of human language has seen an increase of interest in the origins of basic word order.

Givon (1979) observes that SOV basic word order is common among the languages of the world and that many other word orders can be reconstructed back to an SOV stage. He speculates that the first basic word order for

human language was SOV, and other word orders appear to have descended from that order (Givon, 1979). Newmeyer (2000) claims that the earliest human language had rigid SOV order. In comparative linguistics, Gell-Mann and Ruhlen (2011) studied the distribution of the six possible word orders in a sample of 2135 languages, compared them to the putative phylogenetic tree of human languages, and concluded that SOV must have been the word order of the 'ancestral language'.

In sign language linguistics, SOV occupies a special position as well. Studies of newly emerging sign languages (Al Sayyid Bedouin Sign Language and Nicaraguan Sign Language) show a preference for (S)OV word order, despite influences from surrounding languages (Sandler, Meir, Padden, & Aronoff, 2005; Senghas, Coppola, Newport, & Supalla, 1997). In experimental psychology, Goldin-Meadow, So, Özyurek, and Mylander (2008) report an

* Corresponding author. Tel.: +44 753 4319811.
E-mail addresses: Marieke.Schouwstra@ed.ac.uk (M. Schouwstra), H.deSwart@uu.nl (H. de Swart).

experiment in which naive individuals communicated about simple events using only gesture and no speech. They show that speakers of languages with different dominant word orders use the same gesturing order, an order consistent with the word order Subject–Object–Verb (SOV). The authors conclude 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).

It is intriguing that such diverse sources of evidence point in a similar direction: that SOV word order had a special status in the emergence of language. The preference for S initial word order may be traced back to Jackendoff's (2002) observation that in language systems without full grammar, the element that has most control (the agent, or actor) is expressed first in utterances. Dowty (1991) argues that the argument (of a given predicate) that has the most prototypical agent properties is realised syntactically as the subject. However, this do not yet explain why besides SOV, SVO is such a prominent word order in the languages of the world.

Recently, Goldin-Meadow et al.'s experimental methodology was taken up to investigate possible roles of SVO basic word order. It has been suggested that SVO order arises because it is preferred by the computational module in human cognition (Langus & Nespor, 2010), or that SOV/SVO variation comes from communicative pressures: the language user's sensitivity to the possibility of noise corrupting the signal (Gibson et al., 2013). The latter used nonreversible events such as 'rollerskater kicks ball', in which the word order is not essential for obtaining the right interpretation because a ball cannot kick a rollerskater, versus reversible events such as 'fireman kicks girl', in which both nouns could in principle be the agent and word order is essential. They reported an increased usage of SVO order for reversible events. This finding was accounted for by appeal to a general preference to avoid expressing two plausible agents ('fireman' and 'girl') at the same side of the verb. Hall, Mayberry, and Ferreira (2013) report a similar preference for SVO ordering in reversible events, but explain this in terms of cognitive constraints on production: when an event has a human agent and patient (which is typically the case for reversible events), the gesturer will often use the body-as-agent strategy. That is, they act out the agent and patient by 'impersonating' the participants. This strategy leads to confusion when a patient is followed by an action (as it would be the case in SOV order).

Notwithstanding the importance of communicative pressures, and computational constraints on human cognition in shaping natural language, we pursue a different avenue here, and argue that word order is grounded in semantics. In support of this view, we report an improvised gesture experiment. In this task, there is no pre-existing system of linguistic conventions, so people can organise their utterances flexibly. We will show that the choice between SOV and SVO order depends on the meaning of the message to be conveyed. In the experiment we observed the gesturing orders used by naive participants for two kinds of events: extensional and intensional

events. In our experiment, extensional events are instantiated by motion verbs like *throw* or *carry*, also used by Goldin-Meadow et al. (2008). Such verbs are transitive (contain a subject/actor and an object/patient), and involve some action through space. Intensional events (e.g., 'pirate searches guitar,' 'pirate thinks of guitar', but also 'pirate hears guitar' and 'pirate builds guitar'), by contrast, are typically described using intensional verbs (see Forbes, 2010), and for the interpretation of such sentences, the *intension* (meaning) of the direct object is more important than its *extension* (object in the world). Forbes (2010) defines three features that characterise direct objects of intensional verbs: (1) resistance to substitution (i.e., *Mary admires Mark Twain* does not necessarily mean the same as *Mary admires Samuel Clemens*); (2) the possibility of an non-specific reading (such as in the sentence *Mary is looking for a man, but not one in particular*), or (3) existential neutrality (i.e., a sentence like *John is looking for a unicorn* is possible, in which the unicorn does not exist).

We claim that the semantic differences between intensional and extensional verbs form the direct basis of an expected order difference in improvised gesture, by appealing to existing literature. Recall that S-initial word order is expected to be dominant, because of the role of the 'Agent First' principle (cf. Dowty, 1991, Jackendoff, 1992). Goldin-Meadow et al. (2008), who only discuss extensional events, state that O is naturally sequenced before V, because 'entities are cognitively more basic and less relational than actions, which might lead participants to highlight entities involved in an action before focusing on the action itself'. Together these observations allow us to hypothesise that extensional events are gestured in SOV order. Intuitively, an extensional event like 'pirate throws guitar' can be paraphrased as 'You know the pirate? You know the guitar? He throws it.'

Direct objects that are arguments of extensional verbs refer to concrete objects that are identified as existing independently of the event, but intensional verbs take direct objects that are possibly non-specific or non-existent. This makes direct objects in intensional events more abstract and more dependent on the action than those in extensional events, and this is, we hypothesise, a reason to describe them *after* the verb. If one would paraphrase an intensional event like 'pirate thinks of guitar' it is much less natural to use this order to present the information in ('You know the pirate? You know the guitar? He thinks of it.'). Given that the guitar is dependent on the pirate's thoughts, we *need the thought bubble before the object within it*. Thus, for intensional events, we expect the linear order SVO rather than SOV.

We can then use the contrast between extensional and intensional events to test the hypothesis that the meaning of the message conveyed drives the sequencing of gestures in an improvised production task: two different orders (SOV and SVO) are expected for extensional and intensional events, respectively. Moreover, these orders are predicted to occur independently of the structure of the native language of participants, because we take the improvised gesture production to be a communicative system without full grammar, that circumvents the linguistic conventions of the native language. In order to test this, we carried

out experiments with two groups of participants (Dutch and Turkish), whose native language either has SVO or SOV as the dominant word order.

2. Experiment: Improvised gesture production

2.1. Method

2.1.1. Participants

16 Participants (5 male and 11 female) were recruited from Utrecht University and the Utrecht School of the Arts in Utrecht, the Netherlands. All were native speakers of Dutch (an SVO language)[1], and none had any knowledge of a conventional sign language. 19 Participants (10 male and 9 female) were recruited from Bogazici University in Istanbul, Turkey. All were native speakers of Turkish (an SOV language), and none had any knowledge of a conventional sign language. All participants received a small monetary compensation.

2.1.2. Material

The set of items consisted of 20 pictures of extensional events (e.g. 'Pirate throws guitar', 'Princess carries vase'), and 20 pictures of intensional events (e.g. 'Cook thinks of sock', 'Gnome sees tall building'); see Figs. 1 and 2 for examples. Each extensional 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 the participants to really gesture all elements in the picture. All patients (direct objects) in our items were inanimate objects, in order to exclude the animacy/reversibility effects described in Meir, Lifshitz, Ilkbasaran, and Padden (2010), Gibson et al. (2013) and Hall et al. (2013). All pictures had been 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. Two versions of the experiment were created, each consisting of 10 pictures of extensional and 10 pictures of intensional events. Each version contained only one occurrence of each action. The items were presented in random order (and a different order for each participant), and we made sure that two consecutive items in the experiment never contained the same subject or object.

2.1.3. Procedure

Participants were shown pictures of events on a computer screen. They were asked to convey their meanings 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.

Fig. 1. Example item: intensional event. 'Pirate throws guitar'.

Fig. 2. Example item: extensional event. 'Cook thinks of sock'.

Before the actual experiment, participants were shown four practice items. During the practice stage, the experimenter gave feedback about whether or not she had understood which meaning was conveyed. No spoken feedback was given during the experiment. After the gesturing stage, participants were shown the pictures again, and were asked to describe each event using a declarative sentence in their native language.

---

[1] Subordinate clauses in Dutch are SOV ordered. However, our experiment is only concerned with simple events and main clauses, which always have SVO order.
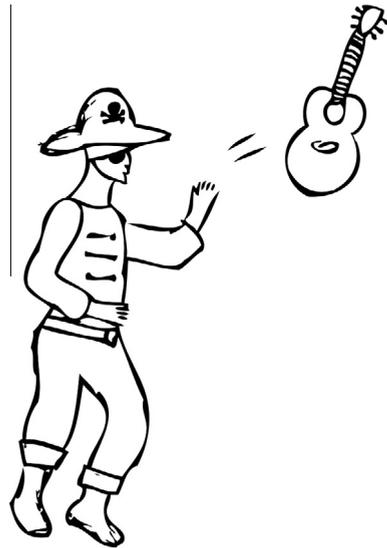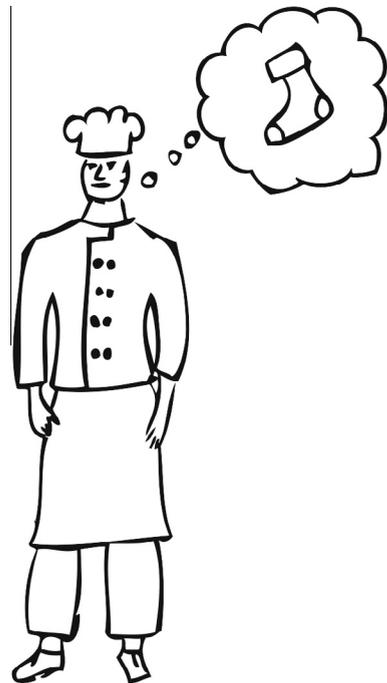
2.2. Results

The video recordings were coded for gesturing order by two independent coders. Occasionally, participants produced gesture strings describing an action that did not match with the intended action on the picture; these were removed (75 of 700 recordings). All gesturing sequences for which there was no consensus in the coding were filtered out as well (39 recordings).[2] Because participants were completely free to choose the gestures to convey the information of each item, they produced SOV strings, SVO strings as well as other orders. The category of 'other' orders (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 were analysed using a repeated measures ANOVA. The within subject factor was Picture-type (intensional or extensional) and the dependent variable was Order (SOV or SVO); the between subjects factors were Version (version 1 or 2) and Language (Turkish and Dutch). We looked at the interaction between Picture type and Order and found this was significant: $F(1,31) = 292.136$, $p = .000$. No significant interaction was found between the main interaction and that of version ($p = .902$) or language ($p = .172$).

Pairwise Bonferroni corrected comparisons reveal (*Figs. 3 and 4*) that among gesture strings of extensional events, the proportion of SOV order was high ($M = .679$, SE = .036), whereas the proportion of SVO order was low ($M = .100$, SE = .035). Among gesture strings of intensional events, the proportion of SVO was high ($M = .516$, SE = .045), whereas the proportion of SOV was low ($M = .108$, SE = .020).

The experiment shows that that the semantic properties of events have an influence on the gesturing order in improvised communication: SOV order is preferred over SVO for extensional (motion) events, and SVO order is preferred over SOV for intensional events. The two kinds of events consisted of an equal number of elements (an actor, a patient and some action), and the same actors and patients appeared in both conditions. The difference between them can only be described in semantic terms: for extensional verbs, the *extension* (reference) of the direct object that is involved is important, whereas for intensional verbs only the *intension* (meaning) of the direct object that is involved is important, which makes the latter much more abstract. Moreover, the fact that our findings were not influenced by the native language of the participants shows that our findings are language-independent, and the gesture task really measures sequencing preferences in a communication system that lacks linguistic conventions.

Fig. 3. Results: extensional events. Mean proportions of SOV and SVO gesturing orders for extensional events. Error bars indicate standard error of the mean.

Fig. 4. Results: intensional events. Mean proportions of SOV and SVO gesturing orders for intensional events. Error bars indicate standard error of the mean.

Our results for extensional events (showing a preference for SOV gesture order) are fully consistent with previous work that adopted the same methodology (Goldin-Meadow et al., 2008; Gibson et al., 2013; Langus & Nespor, 2010). However, the striking difference between the orders preferred for describing intensional and extensional events gives rise to the conclusion that it is not SOV word order as such that is important in the improvised gesture task, but rather the semantic makeup of the message.

3. Discussion

In previous literature it has been claimed that SOV word order must have played a central role in the emergence of language. Although this claim satisfactorily reflects the fact that SOV word order is frequent among the basic word orders of the languages of the world, it does not give us an account of why there is another order that is quite prominent: SVO.

The insight that improvising gesturers use different orders for different semantic classes of events, suggests that we should not think of SOV as a strict principle for

---

[2] The non-matching gesture sequences were selected on the basis of written descriptions of the items, provided by the participants. The coding of gesture sequences was done on the basis of the video recordings; for each recording, the most complete gesture sequence was coded. The sentences and gesture sequences were judged by two independent coders, and the items for which the two coders disagreed were interpreted by the experimenter (who could not see the previous scores).
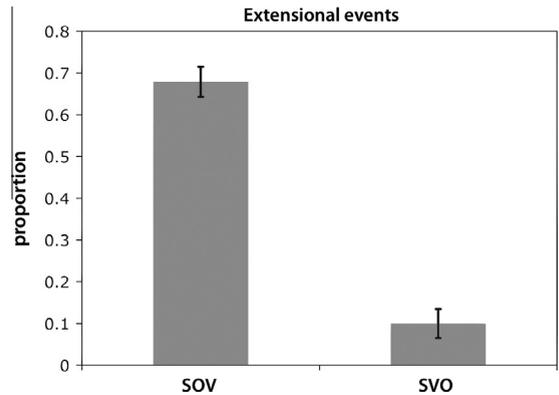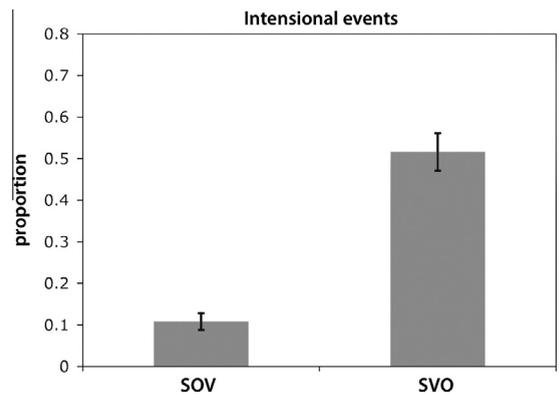
information structuring in emerging communication systems. Rather, in these systems, meaning and structure have more to do with each other than previously thought. Moreover, our results indicate that ordering information in utterances in these systems is quite an active process, rather than simply a reproduction of how the information is represented mentally. In the experiment participants are presented with a holistic image of an event, and they are forced to linearise the information depicted in the image, i.e., they are forced to impose an order on the information. It is only in making the information *public* (in being involved in communication) that ordering plays a role (cf. Slobin, 1996 notion of *thinking before speaking*). As is shown in the experiment, people show language independent systematicity in the ordering they choose. This systematicity is shown to be based on semantic properties.

We are not the first to propose a mode of language in which word order is not strictly determined by syntactic rules—but instead more flexible, semantic principles—play a role. In linguistics, language systems have been reported that arise in situations where speakers cannot learn or use language in a normal way. In these systems, like in the experiment presented here, utterance organisation based on semantics has been observed. An example of such a system is a stage in adult second language learning. In the process of unsupervised second language learning, adults go through a stage that has been characterised as being (1) determined by a small number of semantic/pragmatic organisational principles, (2) largely independent of the source or target language and (3) simple but relatively successful for communication. This stage is called the Basic Variety (Klein & Perdue, 1997). Similar semantic/pragmatic organisational principles were described for, e.g., pidgin languages, and it was claimed that the patterns found in these situations might tell us something about the structure of evolutionarily early language (Jackendoff, 2002).

The experiment described in this paper shares many properties with the settings of pidgin and unsupervised second language learning (Schouwstra, 2012, chapter 4). Just like language contact situations which lack a shared linguistic code, the experimental set-up makes it impossible for the subjects to use their native language to express themselves and forces them to improvise, using whatever means available in their restricted inventory. The fact that there is a conceptual connection between the natural linguistic phenomena and the experimental work presented above, means that two valuable sources of evidence can be combined to paint a more complete picture of the emergence of human language.

Our analysis contrasts improvised language use with fully developed languages like English, and stresses that the former is governed by semantic principles, whereas the latter are organised strictly according to grammatical rules. One could object that this distinction is incorrect, and that semantic organisation is common in full languages as well. Indeed, effects of semantic differences on word order have been observed in full languages to some degree. For instance, the construction of passives is influenced by animacy features of the agent and patient (Branigan, Pickering, & Tanaka, 2008; van Nice & Dietrich, 2003; Ferreira, 1994). Such studies focus on features of

the nominal argument, and the discussion is generally oriented towards subjects. Our investigation is new, because it relates the semantics of the verb to the position of its object. Moreover, it moves away from existing optionality in languages of the world (like active/passive alternations) to variability that is exclusively observed in emergent communication systems that rely on improvised word order rather than grammar.

To conclude, our experimental results show that in improvised communication, 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 linearisation options, and choices are driven by different factors. The pragmatic rule of 'Agent First' favours S-initial word order. The status of objects as independent of the action with extensional verbs makes SOV the preferred word order. Communicative pressures or cognitive constraints may lead to the emergence of SVO word order (Gibson et al., 2012, Hall et al., 2013). Our experiment complements earlier studies by focusing on the semantics of the verb, and shows that intensional events differ from extensional events in favouring SVO word order. The results suggest that the semantic origins of word order should not be overlooked in the debate on language evolution, and semantic organisational principles should be seen as the precursors of syntactic rules.

Acknowledgements

References

Branigan, H. P., Pickering, M. J., & Tanaka, M. (2008). Contributions of animacy to grammatical function assignment and word order during production. *Lingua, 118*(2), 172–189.

Dryer, M. S. (2011). Order of subject, object and verb. In M. S. Dryer & M. Haspelmath (Eds.), *The world atlas of language structures online*. Max Planck Digital Library.

Ferreira, F. (1994). Choice of passive voice is affected by verb type and animacy. *Journal of Memory and Language, 33*(6), 715–736.

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

Gell-Mann, M., & Ruhlen, M. (2011). The origin and evolution of word order. *PNAS, 108*(42).

Gibson, E., Piantadosi, S. T., Brink, K., Bergen, L., Lim, E., & Saxe, R. (2013). A noisy-channel account of crosslinguistic word-order variation. *Psychological Science, 24*(7).

Givon, T. (1979). *On understanding grammar*. New York: Academic Press.

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

Hall, M. L., Mayberry, R. I., & Ferreira, V. S. (2013). Cognitive constraints on constituent order: Evidence from elicited pantomime. *Cognition, 129*(1), 1–17.

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).

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.

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.

Sandler, W., Meir, I., Padden, C., & Aronoff, M. (2005). The emergence of grammar: Systematic structure in a new language. *PNAS, 102*(7).

Schouwstra, M. (2012). Semantic structures, communicative strategies and the emergence of language. LOT dissertation series 312. Utrecht: LOT.

Senghas, A., Coppola, M., Newport, E., & Supalla, T. (1997). Argument structure in Nicaraguan sign language: The emergence of grammatical devices. In E. Hughes, M. Hughes, & A. Greenhill (Eds.), *Proceedings of the Boston university conference on language development* (pp. 550–561). Boston: Cascadilla Press.

Slobin, D. I. (1996). From 'thought and language' to 'thinking for speaking'. In J. Gumperz & S. Levinson (Eds.), *Re- thinking linguistic relativity*. Cambridge University Press.

van Nice, K. Y., & Dietrich, R. (2003). Task sensitivity of animacy effects: Evidence from German picture descriptions. *Linguistics, 41*(5; ISSU 387), 825–850.