IMPROVISED WORD ORDER BIASES ARE NOT MODALITY SPECIFIC: EVIDENCE FROM NON-LINGUISTIC VOCALIZATIONS

KATIE MUDD[*1,2], SIMON KIRBY[1], and MARIEKE SCHOUWSTRA[1]

[*]Corresponding Author: kath.mudd@gmail.com
[1]Centre for Language Evolution, University of Edinburgh, Edinburgh, UK
[2]AI-lab, Vrije Universiteit Brussel, Brussels, Belgium

Languages use different word orders (the order of the Subject, Object and Verb) to organize and convey information. The distribution of word orders documented in the world is uneven, favoring SOV and SVO (Dryer, 2013), which may provide a window into biases shaping language. The visual-manual modality has recently been taken up to study basic word order in the absence of a system of language conventions. In silent gesture experiments, in which hearing participants improvise ways of communicating motion events using only their hands and no speech, SOV order is dominant for speakers of various native languages (Goldin-Meadow et al., 2008). However, participants may switch to SVO word order depending on the semantic properties of the event (Schouwstra & de Swart, 2014; Gibson et al., 2013), or in the presence of a vocabulary of conventional signs (Marno et al., 2015). As these findings appear to be robust across different languages (native SOV and SVO speakers in the aforementioned studies), silent gesture is increasingly used as a way to uncover the cognitive biases playing a role in situations of emerging language.

Thus far, these experiments have been carried out using the manual modality alone, leaving open the question whether the findings generalize to other modalities.[1] Are the word order patterns observed in silent gesture essentially a consequence of the fact that participants improvise in the absence of linguistic rules? Or are they (partly) the result of modality specific production constraints (such as suggested in e.g. Hall et al. 2013)? An answer to this question will have

[1] Schouwstra (2017) observes a structural similarity between silent gesture strings and (spoken) utterances of unsupervised second language learners, but this concerns the placement of temporal information; systematic cross-modal investigations of Basic Word Order have, to our knowledge, not been carried out before.

implications for the generality of the evolutionary claims to be made on the basis of silent gesture results.

Here we use a vocal analog of the silent gesture paradigm to test whether biases previously observed in silent gesture generalize to the vocal modality. We do this by testing the word order preferences for *extensional events* (in which both participants in the action are concrete, e.g., 'the pirate throws the guitar') vs *intensional events* (in which the Patient is unspecific or dependent on the action, e.g., 'the pirate thinks of the guitar'). Schouwstra & de Swart (2014) showed a preference for SOV for extensional events and SVO for intensional events in silent gesture. Here, we tested participants in the vocal modality, to see if this meaning-dependent word order pattern replicates outside the manual modality.

In experiment 1, we asked native SVO participants (N=20) to use non-linguistic vocalizations to describe pictures while sitting on their hands. The stimuli were line drawings that consisted of 32 extensional events (e.g. 'a robot drops a drill') and 32 intensional events (e.g. 'a diver loves a car') that were selected to be relatively easy to describe using non-linguistic vocalizations. In each description, the elements were coded as subject, object or verb. For instance, 'evil-laugh' (denoting a witch) was coded as Subject; 'boom' (denoting dropping of an object) was coded as Verb. Sequences with repeated consecutive constituents were recoded as if having only one iteration of each constituent (SVVO was recoded as SVO). Sequences with non-consecutive repetitions remained as they were (VSVO remained VSVO). Word orders other than SVO or SOV (14% of the total data set) were excluded from statistical analysis.

The findings confirm that constituent order in improvised vocalizations is influenced by the semantic properties of the event: a greater proportion of SOV descriptions are given for extensional events than intensional events, and a greater proportion of SVO descriptions are given for intensional events than extensional events (see Figure 1). Event type predicted the proportion of SVO produced (b = 1.965, SE = .737, $p < .01$), mirroring what has been found in silent gesture.

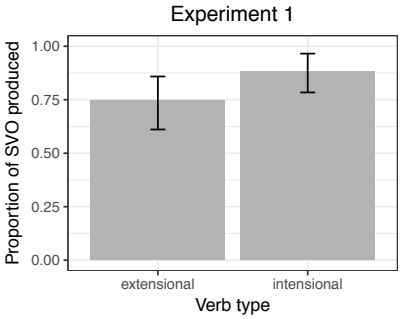

Figure 1. Mean proportions of SVO produced for extensional and intensional events in experiment 1 out of all SVO and SOV produced. Error bars represent 95% confidence intervals.

Strikingly, the results from the vocal modality follow what has been found in the manual modality, namely that semantic properties of the event significantly affect constituent order. However, the results of experiment 1 differ from previous findings, where the preference for SOV for extensional events is much greater, and SVO only becomes dominant for the intensional events (Schouwstra & de Swart, 2014). A possible explanation for the difference in results is that in the present study the vocal channel is used for both language and the improvised vocalizations, and thus may be more susceptible to interference from the native language of participants (SVO in the present study). Despite this possible influence, the vocalizations of participants in the present study were conditioned by meaning, and we thus replicated the basic finding of Schouwstra & de Swart (2014).

In experiment 2, we build on work by Marno et al. (2015) who show an increase in SVO order in silent gesture if participants (native SOV and SVO) are trained on individual lexical items (e.g. 'a girl') prior to complex meanings, akin to the extensional events in the present study (e.g. 'a girl throwing a fish'; Marno et al., 2015). We asked native SVO participants (N=20) to produce vocalizations for individual items first, before going on to complete scenes. Using the same coding scheme as for experiment 1, word orders other than SVO or SOV (9% of the total data set) were excluded from statistical analysis.

Again, our results in the vocal modality reflect those in the manual modality: a higher proportion of SVO is produced when participants first describe individual items before continuing to full scenes. In experiment 2, SVO accounts for a majority of extensional events (87%) and intensional events (99%) while SOV accounts for a minority of extensional events (13%) and intensional events (1%). Overall, more SVO is produced for both event types in experiment 2 than in experiment 1; in a model combining experiment 1 and 2, we find that experiment (b = 3.998, SE = 1.465, $p < .01$) is a stronger predictor of producing

SVO than verb type (b = 1.693, SE = .714, $p < .05$); thus, the high incidence of SVO in experiment 2 is more strongly predicted by the access to the lexicon than the meaning of the event. Following the results of Marno et al. (2015) in the manual modality, the lexicon triggers more SVO, showing similar results in the vocal and manual modalities.

The silent gesture paradigm has been used to make claims about features of language emergence, in the manual and vocal modalities, but could these findings be specific to the manual modality? In the present study, we focused on 2 results from the silent gesture paradigm, and replicated both of the main findings in the vocal modality. In the vocal modality, more SVO is produced than in the manual modality, which may be attributed to interference from the participant's native language, but the overall finding is the same: word order is conditioned by event type in improvisation. These findings provide the first evidence that gestural and vocal improvisation yield similar results for basic word order, demonstrating that the effects observed are modality independent.

References

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), 1079-1088.

Goldin-Meadow, S., So, W. C., Özyürek, A., & Mylander, C. (2008). The natural order of events: How speakers of different languages represent events nonverbally. *Proceedings of the National Academy of Sciences*, 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.

Marno, H., Langus, A., Omidbeigi, M., Asaadi, S., Seyed-Allaei, S., & Nespor, M. (2015). A new perspective on word order preferences: the availability of a lexicon triggers the use of SVO word order. *Frontiers in psychology*, *6*.

Schouwstra, M. (2017). Temporal Structure in Emerging Language: From Natural Data to Silent Gesture. *Cognitive science*, *41*(S4), 928-940.

Schouwstra, M., and de Swart, H. (2014). The semantic origins of word order. *Cognition* 131, 431–436. doi: 10.1016/j.cognition.2014.03.004