INTERPRETING PATIENT ANIMACY IN SILENT GESTURE

ROSS TOWNS[1*], MARIEKE SCHOUWSTRA[2]

[*]Corresponding Author: s1667434@sms.ed.ac.uk
[1, 2]Centre for Language Evolution, University of Edinburgh, Edinburgh, United Kingdom

A key finding of studies in the silent gesture paradigm is that people prefer SOV order when describing events in which an action's agent (corresponding to a verb's subject) is animate and the patient (object) is inanimate (Goldin-Meadow, So, Ozyürek, & Mylander, 2008). For so-called reversible events, however, in which agent and patient are both animate (and whose thematic roles may therefore plausibly be reversed), they prefer alternatives like SVO or OSV (Futrell et al., 2015; Gibson et al., 2013; Hall, Mayberry, & Ferreira, 2013; Meir, Lifshitz, İlkbasaran, & Padden, 2010).

Three competing explanations for this have been put forward. Under the so-called *ambiguity* hypothesis, participants avoid SOV due to the potential ambiguity caused by positioning two possible subjects before the verb (Meir et al., 2010). The *noisy-channel* account cites the susceptibility of SOV utterances to information loss caused by noise, to which SVO and related orders present a more robust solution (Gibson et al., 2013). Finally, the *role-conflict* account cites the tendency of gesturers to physically embody the roles of animate referents, thereby leading to potential confusion between roles if patients are gestured before actions (Hall et al., 2013).

Here we present an experiment designed to identify whether a relationship between patient animacy and gesture order is apparent in interpretation as well as production. We asked participants to interpret ambiguous gesture sequences, predicting that SVO sequences would be more likely to be interpreted as describing reversible events than would SOV sequences, and vice versa. However, our findings did not support this: gesture order did not seem to influence interpretation of reversible/non-reversible events. This result led us to carefully reconsider the mechanisms behind previous silent gesture production findings. We suggest that our null result is consistent with an account similar to the role-conflict hypothesis described above, in which apparent effects of

reversibility on word order are due to the specific affordances of the gestural modality, and are of limited generalisability to spoken language.

Our study builds on previous investigations of word order alternations in silent gesture in another different semantic domain: that of *intensional* vs *extensional* events (Schouwstra, 2012). By showing participants videos with ambiguous action/verb gestures, Schouwstra (2012; see also Thompson, Schouwstra & de Swart, 2016) found that people tended to interpret SOV gestures as describing extensional events (i.e. actions performed on an entity in the physical world, for example *girl sleeps on book*) but that SVO gestures were interpreted as describing intensional events (i.e. for example mental actions such as *girl dreams of book)*. Participants' biases in interpretation were thus consistent with, though somewhat weaker than, a bias found in gesture production for intensional and extensional events (Schouwstra & de Swart, 2014). A Bayesian computational model fitted to data from the interpretation study was also found to align well with data from production (Thompson, Schouwstra & de Swart, 2016), suggesting that production and interpretation are subject to related cognitive biases. Based on these results, we reasoned that we might find a similar correspondence between production and interpretation in the domain of reversibility.

Basing our experimental design on Schouwstra's (2012) interpretation study, we showed participants ambiguous gesture videos in 12 forced-choice trials. For each trial, we recorded 2 videos of gesture sequences, both consisting of the same gestures recorded in separate shots and concatenated either in SOV or SVO order using video editing software. Crucially, gestures for patients/objects were designed to be ambiguous as to whether they represented an inanimate or animate referent, for example a *guitar-playing* gesture, which might represent either a guitar or a guitarist. Participants watched SOV videos in 6 trials and SVO videos in 6 trials (choice of video randomized between trials), and signaled their interpretation by choosing from an array of two line-drawings, one depicting a reversible event and another depicting a non-reversible event (e.g. *pirate strokes guitarist* or *pirate strokes guitar).* We conducted two runs of the experiment: once using online participants recruited via the CrowdFlower crowdsourcing platform (N = 51, native English speakers) and once with participants in the lab (N = 20, native English speakers) so as to corroborate the results of the online study.

Contrary to our expectations, we found in both runs that SVO videos were marginally less likely to be interpreted as reversible than SOV videos (see Fig. 1). To account for possible variation among participants and trial items, we fit data from each run, respectively, to mixed-effects logistic regressions (specifying by-participant and by-item random slopes and intercepts) predicting

reversible interpretation by SVO gesture order, and used likelihood-ratio tests to compare them with null models featuring random effects only. Inclusion of gesture order
as a predictor was not found to improve model fit for either the online experiment ($\chi 2$ = 1.11, p = .29) or the laboratory experiment ($\chi 2$ = .008, p = .93).

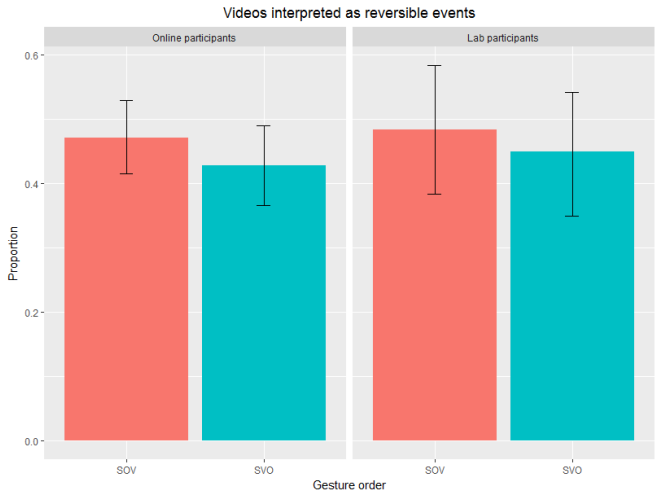

Figure 1. Mean proportions of gesture videos interpreted as reversible events in the online experiment (left) and the lab experiment (right). Error bars represent bootstrapped 95% confidence intervals.

Why, then, should biases in interpretation correspond to production with respect to intensionality, but not to reversibility as in the present study? It may simply be that the stimuli for our experiment simply were not ambiguous enough for any effect to take hold. However, another possibility points to qualitative differences between the mechanisms underpinning production biases in each domain. With respect to intensionality, findings from production and interpretation suggest a bias toward positioning elements in order of abstractness or cognitive accessibility (Goldin-Meadow et al., 2008; Schouwstra, 2012), or an iconic linear sequence in which object-final order mirrors the causal relations between an intensional object on its verb, for example between a thought and the action of thinking (Christensen, Fusaroli, & Tylén, 2016; Schouwstra & de Swart, 2014). Both accounts are thus predicated on properties of SOV and SVO word order *per se*, and can be considered apart from gesture production.

The same may be said of reversibility effects under the ambiguity and noisy channel hypotheses, leading to the expectation that participants might exhibit a bias toward interpreting SVO gestures as reversible. By contrast, the role conflict account is rooted in gesture production, and is therefore less likely to suggest a corresponding bias in interpretation. In support of this production-specific account, Hall and colleagues cite the results of their interpretation study, which found no effect of order on interpretation, either in terms of the purported ambiguity of SOV order or in participants' ratings of SOV and SVO sequences as either more or less appropriate descriptions of visual stimuli (Hall, Ahn, Mayberry, & Ferreira, 2015).

While remaining mindful of the dangers inherent in drawing conclusions from a null result, we suggest that our findings are consistent with an account of reversibility/patient animacy effects on gesture production as being rooted in the gestural modality itself. If this is correct, the extent to which silent gesture findings, as pertaining to reversibility/patient animacy, can be generalised to spoken language is debatable. Hall et al. (2013) suggest that avoidance of role-conflict may shape word-order in spoken language via perspective-taking (MacWhinney, 1977), but this is offered as a tentative suggestion without strong empirical support.

Our interpretation of the present findings does, however, generate two testable predictions, namely that the relationship between intensionality and constituent order in gesture should also be found in both production and interpretation in a different modality, for example visual symbols (Vastenius, van de Weijer, & Zlatev, 2016), whereas effects of reversibility/patient animacy will be confined to production in the gestural modality only. Our results also, in our view, demonstrate the need for caution against interpreting possibly modality-specific effects as general properties of cognition, and for variety in the use of alternative modalities as experimental paradigms.

References

Christensen, P., Fusaroli, R., & Tylén, K. (2016). Environmental constraints shaping constituent order in emerging communication systems: Structural iconicity, interactive alignment and conventionalization. *Cognition*, *146*, 67–80.

Futrell, R., Hickey, T., Lee, A., Lim, E., Luchkina, E., & Gibson, E. (2015). Cross-linguistic gestures reflect typological universals: A subject-initial, verb-final bias in speakers of diverse languages. *Cognition*, *136*, 215–221.

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., Ozyürek, A., & Mylander, C. (2008). The natural order of events: how speakers of different languages represent events nonverbally. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(27), 9163–9168.

Hall, M. L., Ahn, Y. D., Mayberry, R. I., & Ferreira, V. S. (2015). Production and comprehension show divergent constituent order preferences: Evidence from elicited pantomime. *Journal of Memory and Language*, *81*, 16–33.

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

MacWhinney, B. (1977). Starting Points. *Language*, *53*(1), 152-168.

Meir, I., Lifshitz, A., İlkbasaran, D., & Padden, C. (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 (EvoLang 8)*, 455–456.

Schouwstra, M. (2012). *Semantic Structures, Communicative Strategies and the Emergence of Language*. Utrecht: LOT

Schouwstra, M., & de Swart, H. (2014). The semantic origins of word order. *Cognition*, *131*(3), 431–436.

Thompson, B., Schouwstra, M., & de Swart, H. (2016). Interpreting silent gesture. In S. G. Roberts, C. Cuskley, L. McCrohon, L. Barceló-Coblijn, O. Feher, & T. Verhoef (Eds.), *The Evolution of Language: Proceedings of the 11th International Conference (EVOLANG11).*

Vastenius, A., van de Weijer, J., & Zlatev, J. (2016). The influence of native language word order and cognitive biases in pictorial event representations. *Cognitive Semiotics*, *9*(1), 45.