Evidence for argument structure in the form of pantomime

The emerging field of 'Super Semantics' asks whether semantic theory may extend to non-standard objects of study, such as gesture [1]. Gestures may convey pre/cosuppositional, at-issue, or supplemental information, depending on their relation to speech. When gesture takes on the full communicative load (*pantomime*), further language-like properties emerge. Pantomime, thus, may contain the raw material on which sign languages are built [2].

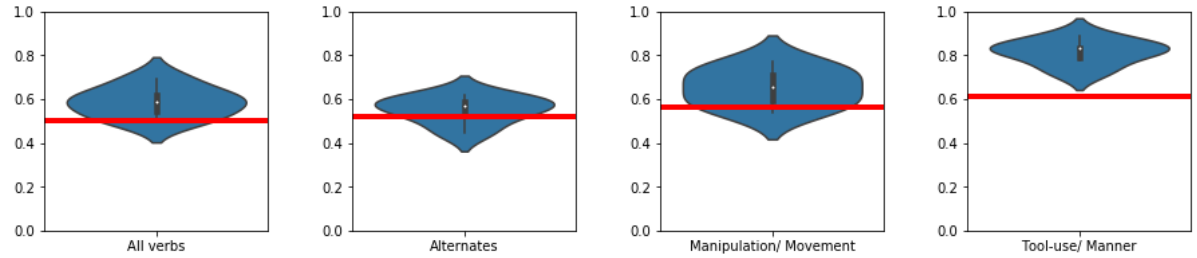
One area of intense investigation is the emergence of transitivity distinctions in pantomime: gesturers, like signers, manipulate handshape to differentiate transitive and intransitive events [3, 4]. However, work to date is not readily generalizable as it concerns (a) relatively few event types, (b) just two handshape features, and (c) whether the proportion of these features is higher in one class of predicate or the other. We argue that this simultaneously underestimates the total information available in the signal while overestimating the importance of a selected few. Here, we consider a broader range of events and handshape features. In doing so, we are able to make inferences about the internal, hierarchical structuring of pantomime.

Method: We elicited single-gesture utterances from 6 participants. Pantomimes were produced in response to vignettes of 72 unique events that involved the manipulation (transitive) or movement (intransitive) of a variety of objects. We annotated these 432 (6 * 72) productions for 6 handshape parameters individually linked to transitivity marking in sign languages. We trained linear support vector classifiers to predict whether each pantomime is in/transitive, using a 6-fold leave-one-out paradigm: The data were split into 6 partitions, trained on 5 of the partitions and tested on the 6th, producing an accuracy score. This was done in a round-robin fashion such that each partition was the test set once. We computed mean accuracy and compared it against chance using the probability mass function of the binomial distribution. To assess the degree semantic class affects handshape production, the analysis was run on 3 subsets of the data: *Causative/ inchoative (alternating) verbs* (n = 174), *Verbs of manipulation and movement* (n = 234), and *Verbs of tool-use and manner* (n = 108). (The first two sets are disjoint; the last set is a proper subset of manipulation/ movement verbs.) Finally, to assess feature importance, we averaged the model weights for each predictor in each analysis.

Results: Classifiers trained on all verbs in the dataset achieved a mean accuracy of 59% ($p < 0.001$). Performance on alternating predicates was poor (54%, n.s.), but performance on verbs of manipulation/ movement and tool-use/ manner was good (67%, $p = 0.003$; and 82%, $p < 0.0001$, respectively). See Figure 1. Among manipulation/ movement and tool-use/ manner verbs, four predictors were consistently identified as relevant to transitivity marking (Table 1). All predictors are associated with *transitive* items (no predictor crossed 0 across Folds), showing that transitives are overtly and exhaustively marked.

Interpretation: The analyses show that transitivity information is present in the form of pantomimed actions. Further, pantomimes do not seem to form a monolithic class, but vary predictably in their form-meaning correspondences, building on results in [3, 4]. The analysis also suggests that certain, narrowly defined semantic categories may enjoy a reliable coding preference that is not available to broader classes.

Finally, the analysis of the predictors indicate that transitive pantomimes are phonetically more complex than intransitives and suggests that their syntax is similarly more complex (e.g., involve more projections). We develop an analysis that identifies each predictor as the spell-out of specific nodes of hierarchical structure (e.g., the curvature of the hand maps to the internal argument of the predicate). We use this analysis to explain the why classifiers trained on different subsets of data had very different outcomes.



(a) (b) (c) (d)

Figure 1: Violin plots showing distribution of classifier accuracies (0.0 = 0% accurate, 1.0 = 100% accurate). Red line represents blind baseline (chance).

Analysis/ Features	2-handed?	Finger Complexity	Flexion	Flexion (NSF)
All verbs	1.1995	0.1928	0.3761	–
Alternates	1.6667	–	–	-0.3442
Manip/Mvmt	0.6691	0.3559	0.7668	0.2083
Tool/Manner	0.4604	0.4444	1.0999	0.0723

Table 1: Average model coefficients for the four best predictors. Positive values correspond with 'transitive' labels. '2-handed?' = whether the production involved one or two hands (categorical); 'Finger complexity' = a measure of ease of articulation (1 to 4); 'Flexion' = degree of curvature of the selected (profiled) finger (1 to 7); 'Flexion (NSF)' = degree of curvature of the non-selected (backgrounded) fingers (-1 to 1). Note that in the analysis of all verbs and of alternate verbs, not all predictors were consistently selected as being most informative.

References

[1] Philippe Schlenker. What is super semantics?*. *Philosophical Perspectives*, 32(1):365–453, 2018.

[2] Ronnie B Wilbur and Evguenia Malaia. Contributions of sign language research to gesture understanding: What can multimodal computational systems learn from sign language research. *International journal of semantic computing*, 2(01):5–19, 2008.

[3] Diane Brentari, Marie Coppola, Laura Mazzoni, and Susan Goldin-Meadow. When does a system become phonological? Handshape production in gesturers, signers, and homesigners. *Natural Language & Linguistic Theory*, 30(1):1–31, 2012.

[4] Diane Brentari, Marie Coppola, Pyeong Whan Cho, and Ann Senghas. Handshape complexity as a precursor to phonology: variation, emergence, and acquisition. *Language Acquisition*, 24(4):283–306, 2017.