Structural iconicity predicts word order in improvised gestures

Introduction. Although the world's languages vary in terms of the order of the Subject, Verb, and Object (SVO, SOV, VSO, VOS, OVS, OSV), the observation that most languages have either SOV or SVO order [1] has led to hypotheses that such orders reflect some underlying/ default cognitive representation of event structure [2]. This research uses the "silent gesture paradigm", in which participants are exposed to non-linguistic depictions of events and are asked to communicate what they think they saw using improvised gesture/pantomime. Using silent gestures allows for non-linguistic communication, thereby revealing how events (i.e., verb) and their participants (i.e., arguments) are linearized *without* the word order restrictions of the speaker's L1; it is proposed that such gestures reflect an earlier stage along the pathway of language evolution to the present day [3].

The underpinnings of word order. Results from studies of using silent gesture show that speakers converge on a SV order for intransitive events a SOV for transitive ones (regardless of the order of participants' L1; [2]). However, Schouwstra & de Swart [4] found that **intensional** transitive verbs (e.g., *seek*, *dream of*) are, in fact, predominantly gestured with a Verb-Object (VO) order (again, regardless of L1; [2] used **extenstional** verbs only). Because intensional predicates involve referentially opaque objects or objects that may not exist [5], [4] conclude that VO orders arise due to the object of an intensional verb being more abstract: the interpretation of the object is dependent on interpretation of the verb in a way that is *not* the case with extensional predicates. [4]'s reasoning was later challenged by Christensen et al. [6], who noted a preference for VO gesturing order when the object is **created** by the verb itself (e.g., *ballerina painting airplane*). They attribute this order to temporality: Because created objects only exist *after* the creation event began, they temporally follow the verb (hence VO gesture order, where linearization of elements reflects temporal sequencing); this is termed "**structural iconicity**". Taken together, the findings of [4] and [6] provide two possible explanations for VO preference in transitive verbs of creation: Either (i) verbs of creation are *themselves* intensional, and VO is due to object abstractness [4], or (ii) word order biases are fundamentally driven by the temporal sequencing of an event (i.e., structural iconicity) [6]. This second explanation adopts the analysis of den Dikken et al. [7], according to which intensional transitive verbs take clauses as their arguments, the embedded clause in (1) thus contains a hidden predicate *have* (see also [5]).

(1)   a.   I want a bike   =   b.   I want [PRO to have a bike]

This would mean that the VO order found in gestured sentences with intensional verbs can be subsumed by the structural iconicity hypothesis: The matrix clause of (1a) would have a Verb-Clause order because the situation time of the wanting situation starts before the situation time of having a bike. Then, based on [2], the embedded clause would be SOV (i.e., PRO-Object-Verb), yielding the SVO order in (2):

(2)   I + want + [~~PRO~~ + bike + ~~have~~]   ➔   I + want + bike   (SVO)

Experiment. We put the structural iconicity hypothesis to the test in a novel gesture experiment in which our focus is on syntactic **subjects** of lexical verbs**.** Importantly, because such subjects can *never* be intensional (only **objects** can), the intensionality hypothesis [4] predicts that subjects should not appear postverbally in a silent gesture paradigm (i.e., no VS), but objects alone should (i.e., VO). We took 8 pairs of intransitive events where the subject (i) exists before the event (e.g., *leaf burns away*) vs. (ii) gets created by the event (e.g., *leaf grows out*). We then

created transitive variants, where the noun is now in *object* position (pre-existing: *unicorn eats leaf*, created: *unicorn spits out leaf*; [6]), yielding a 2x2 (TRANSITIVITY x EXISTENCE; within-subjs) design. 32 participants (none had knowledge of sign language) saw 32 animations of events (Table 1) on a screen and were asked to relay them to the experimenter using silent gestures. All were non-reversible events shown as black & white hand-drawn animations. All participants were proficient in English; some speakers were proficient in other SVO languages (e.g., Spanish), as well as SOV (e.g., Hindi-Urdu), and VSO (e.g., Tagalog). Crucially, no participant spoke a language with a word order that distinguishes between pre-existence vs. creation. After the gesturing round, participants saw the same events, and relayed them to the experimenter orally, in English.

Results. 219 trials were excluded due to missing a verb gesture or being ambiguous/repetitive. We examined the 805 remaining trials, asking first about <u>completeness</u> of the gesturing sequence. We coded trials as "complete" if they consisted of minimally a gesture for the verb plus a gesture for the transitive object (for the transitive conditions) or intransitive subject (for the intransitive conditions); collectively, we call this the "Preferred Argument" ("PrefArg") following [8]. A gesture sequence comprising the verb only, or fusion of PrefArg and verb, was coded as "incomplete/incorporated". The results are shown in Figure 1. A 2x2 logistic mixed-effects regression (crossed random effects for participants and items [9], with maximal random effects structure as per [10]) revealed, most importantly, an interaction of TRANSITIVITY*EXISTENCE ($z =$ -4.9; $p < .0001$), where completeness was *lower* for <u>created</u> subjects than for <u>pre-existing</u> ones (there was no contrast for objects, however). This invokes cross-linguistic patterns of noun incorporation, where subjects can *only* incorporate if they are non-agents (this characterizes all <u>created</u> subjects; but only some pre-existing ones), whereas object incorporation is not generally likewise dependent on object existence [11]. Our second analysis focused on complete sequences only ($n = 527$), asking for each one, whether the PrefArg was gestured *pre*verbally (yielding SV or OV), or not (here, the PrefArg was *post*verbal: VS or VO). The results are shown in Figure 2. There were main effects of both TRANSITIVITY (PrefArgs were preverbal more when they were subjects than objects; $z = 5.7$, $p < .0001$) and EXISTENCE (created PrefArgs were preverbal less than pre-existing ones; $z = -8.3$, $p < .0001$); there was no interaction ($p = .49$). For the oral round of the experiment, SV or SVO order was used in $> 99\%$ of trials.

Discussion. We found that **subjects** (as well as objects, as in [6]) are less likely to be gestured *pre*verbally when they get created by the event (i.e., postverbal order emerges). Because verbal intensionality pertains to *objects* only (i.e., not to subjects), we conclude that intensionality/ abstractness is not the underlying trigger for the VO gesture ordering preference (as per [4]). Instead, we maintain that these types of ordering preferences – and, ultimately therefore, word order preferences across languages – arise primarily from structural iconicity, where the cognitive default order reflects the temporal sequencing of an event. Importantly, this proposal accounts for a hereto forth unexplained typological observation; namely, that the majority of the world's languages (almost 90%, [1]) exhibit SV order. We reason that this reflects a combination of **iconicity** [6] plus **grammaticalization** [12]: For *most* verbs, the subject pre-exists (i.e., verbs with created subjects are relatively less common/frequent), and this more frequent pattern is what gets regularized syntactically. Beyond subjecthood, we hypothesize that typologically common word order patterns, generally, are shaped by the temporal ordering of events.
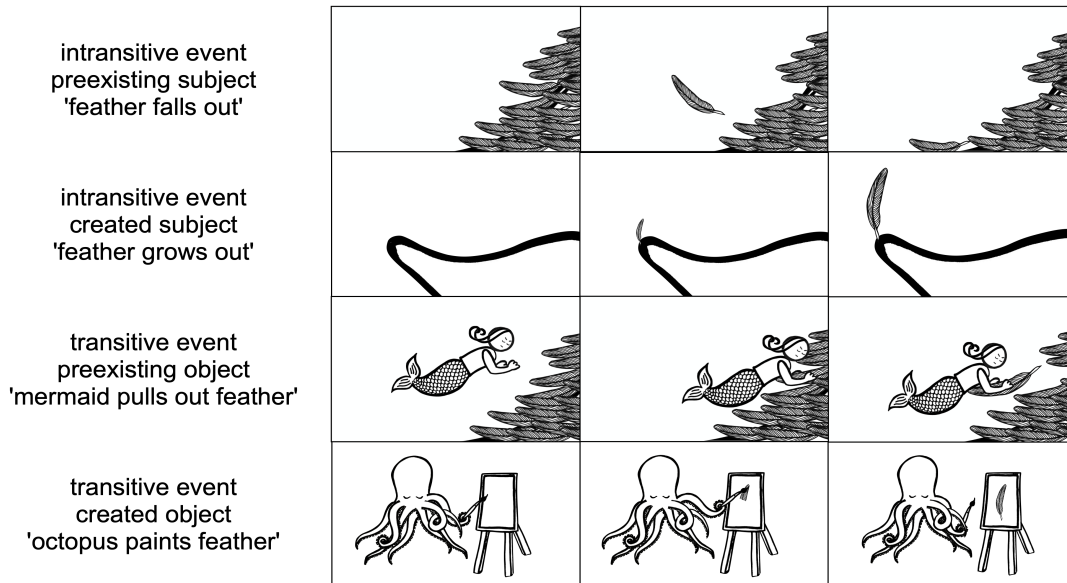
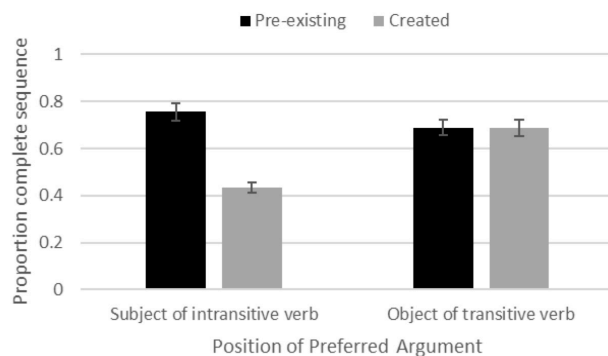Table 1. Captions (left-to-right) from video stimuli

intransitive event preexisting subject 'feather falls out'			
intransitive event created subject 'feather grows out'			
transitive event preexisting object 'mermaid pulls out feather'			
transitive event created object 'octopus paints feather'			

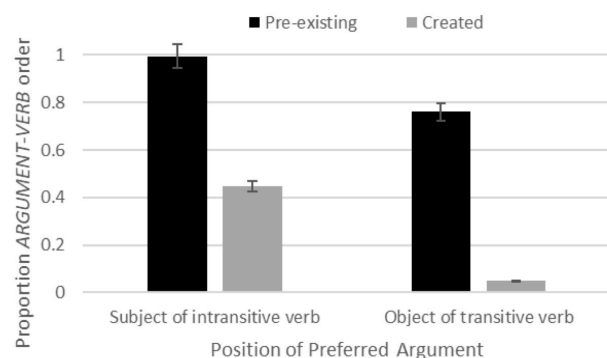Figure 1. Proportions of complete gesture sequences

Figure 2. Proportions of pre-verbal argument gestures (for complete sequences) by condition.

References.

[1] **Dryer (2013).** *WALS.*

[2] **Goldin-Meadow et al. (2008).** The natural order of events:... *PNAS.*

[3] **Tomasello (2008).** *Origins of human communication*. MIT Press.

[4] **Schouwstra & de Swart. (2014).** The semantic origins of word order. *Cognition.*

[5] **Forbes (2020).** Intensional transitive verbs. In *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta.

[6] **Christensen et al. (2016).** Environmental constraints shaping constituent order in emerging communication systems: Structural iconicity, ... *Cognition.*

[7] **den Dikken et al. (2018).** Intensional transitive verbs and abstract clausal complementation. In *Non-propositional intentionality*, OUP.

[8] **Du Bois (1987).** The discourse basis of ergativity. *Language.*

[9] **Baayen, et al. (2008).** Mixed-effects modeling with crossed random effects... *J Mem Lang.*

[10] **Barr et al. (2013).** Random effects structure for confirmatory hypothesis testing... *J Mem Lang.*

[11] **Massam (2009).** Noun Incorporation: Essentials and Extensions. *Lang. Linguist. Compass.*

[12] **Hawkins (2004).** *Efficiency and Complexity in Grammars*. OUP.