A noisy-channel account of crosslinguistic word order variation

Edward Gibson, Steven T. Piantadosi, Kimberly Brink,

Leon Bergen, Eunice Lim & Rebecca Saxe

Brain and Cognitive Sciences Department, Massachusetts Institute of Technology

Send correspondence to:

Edward Gibson, 46-3035, MIT, Cambridge, MA 02139; Email: egibson@mit.edu

Manuscript dated August 27, 2012

4025 words

Abstract

The distribution of word orders across languages is highly non-uniform, with Subject-Verb-Object (SVO) and Subject-Object-Verb (SOV) orders being prevalent. Recent work suggests that SOV may be the default order in human language. Why then is SVO order so common? We hypothesize that SOV/SVO variation can be explained by language users' sensitivity to the possibility of noise corrupting the linguistic signal. In particular, the noisy-channel hypothesis predicts a shift from the default SOV order to SVO order for semantically reversible events, where potential ambiguity arises in SOV order because two plausible agents appear on the same side of the verb. We evaluated and found support for this prediction in three languages (English, Japanese, Korean) using a gesture-production task, which reflects word-order preferences largely independent of native language. Other crosslinguistic variation patterns (e.g., prevalence of case-marking in SOV languages, and its lack in SVO languages) also straightforwardly follow from the noisy-channel hypothesis.

Introduction

It has long been known that the possible orders of the basic units of a clause – the subject (S), verb (V) and object (O) – are highly non-uniformly distributed across languages:

(1) In 1017/1056 (96.3%) of the languages with a dominant word order, subjects precede objects (Dryer, 2005; cf. Greenberg, 1963), and most of the exceptions to this generalization have been argued to be spurious (Dryer, 2002).

(2) Two word orders – SVO (41.2%; e.g., English: *the boy (S) kicks (V) the ball (O)*) and SOV (47.1%; e.g., Japanese: *shonen-ga (“boy”) boru-o (“ball”) kero (“kicks”)*) – are much more prevalent than the third subject-before-object word order, VSO (8.0%).

A plausible explanation for the first generalization is that people tend to construct their utterances from the perspective of agents rather than patients (e.g., MacWhinney, 1977). However, until now, no explanation has been provided for the crosslinguistic prevalence of SOV and SVO word orders. Indeed, the inability of functionalist approaches to explain this distributional pattern (Hockett, 1960; Pinker & Bloom, 1990; Haspelmath, 1999; Hawkins, 2004) has contributed to the argument that grammars are independent of communicative and performance factors and are determined by an innate Universal Grammar (Chomsky, 1986; Baker, 2001).

Here we present a communication-based explanation for the prevalence of SOV and SVO orders and for the crosslinguistic OV/VO variation, building on recent communicative accounts of similarly unexplained linguistic features such as ambiguity

(Piantadosi, Tily, & Gibson 2012). The starting point for this account is the observation that SOV word order appears to be the "default" word order in human language (Givon, 1979; Newmeyer, 2000a, 2000b; Gell-Mann & Ruhlen, 2011). We can break down this preference for SOV into: (a) a preference for subjects to precede objects (explained above); and (b) a preference for the verb to appear clause-finally. With respect to the latter preference, two sources of evidence suggest that there is an initial bias to place the verb after its arguments when developing a communication system. First, two sign languages that were created independently from home-sign systems are verb-final (either SOV or OSV): Nicaraguan Sign Language (Senghas et al., 1997), and Al-Sayyid Bedouin Sign Language (Sandler et al., 2005). Furthermore, Goldin-Meadow et al. (2008) have recently observed that a verb-final order (specifically, SOV) is preferred in a task where participants gesture event meanings, which essentially requires developing a new communication code. Importantly, SOV gesture-production occurs not only for speakers of SOV languages, such as Turkish, but also for speakers of SVO languages, such as English, Chinese, Spanish (Goldin-Meadow et al., 2008), and Italian (Langus & Nespor, 2010), which suggests that this task reflects word order preferences somewhat independent of the person's native language.[1]

If the SOV word order is the default word order in human language, why then is SVO order so prevalent? In other words, why don't all, or most, languages use SOV

[1] The preference for clause-final verb placement can plausibly be explained in terms of the crosslinguistic bias to present old information before new information (Paul, 1880; Jackendoff, 1972): the arguments of a verb are typically old information (already present in the context), and should therefore precede the new information, the verbal predicate. Consistent with this explanation, Schouwstra et al. (2011) demonstrated that people tend to gesture extensional verbs like "kick" and "push" clause-finally, but intensional verbs like "create" (whose objects are new information) clause-medially. Extensional verbs plausibly drive the word order within a language because they appear to be easier for children to acquire (e.g., the average age of acquisition of the verbs in Schouwstra et al. (2011) was 3.99 for the extensional verbs and 5.46 for the intensional verbs (Kuperman, Stadhagen-Gonzales & Brysbaert, 2012)).

order? We propose that SVO order arises crosslinguistically from SOV order due to communicative / memory pressures that can sometimes outweigh the default SOV bias. In particular, building on Shannon's (1949) communication theory, we assume that language comprehension and production operate via a *noisy channel* (Smith, 1969; Aylett & Turk, 2004; Levy, 2008; Levy et al., 2009; Jaeger, 2010; Gibson & Bergen, 2012). A speaker wishes to convey a meaning m and chooses an utterance u to do so. This utterance is conveyed across a channel that may corrupt u in some way, resulting in a received utterance $\tilde{u}$. The noise may result from errors on the side of the producer, external noise, or errors on the side of the listener. The listener must use $\tilde{u}$ to determine the intended meaning m. The best strategy for a speaker is thus to choose an utterance u that will maximize the listener's ability to recover the meaning given the noise process.

One way to evaluate the noisy-channel hypothesis is to compare sentences where the ease of recovering the intended meaning is, versus is not, affected by the order of the elements. Consider, for example, non-reversible sentences: e.g., "girl kicks ball". The word order has little effect on how easily the meaning can be recovered, because the subject (agent) and object (patient) are clear from the semantics – a ball cannot kick a girl. In these situations, people should adhere to the default order, SOV. However, for semantically reversible sentences ("girl kicks boy"), noise may lead to confusion about which is the subject and which is the object in the SOV word order. Gibson & Bergen (2012) provide evidence that English speakers assume a noise process where deletions are most likely, and insertions and transpositions are less likely. If either noun in the SOV sentence like "girl boy kicks" is lost due to noise (resulting in "girl kicks" or "boy kicks"), then the thematic role of the remaining noun phrase is ambiguous: the solitary noun could either be agent or patient. Critically, if SVO word order is used instead ("girl kicks boy"), a deletion will not change how the remaining noun phrase is interpreted: "girl kicks" will allow the listener to recover the meaning of the girl kicking someone/something, and "kicks boy" will allow the listener to recover the meaning of the boy being kicked. In other words, the positions of the noun phrases with respect to the verb can provide a cue about whether the noun is the subject or the object.

Note that although the noisy-channel hypothesis is motivated by a communicative theory, it need not be restricted to situations where we communicate with other people: it applies even if there is only one individual, who is encoding an event meaning for him/herself. Under the noisy-channel hypothesis, the individual will choose a representation that maximizes meaning recoverability (Brady, Konkle, & Alvarez 2009). Indeed, Goldin-Meadow et al. (2008) observed the preference for SOV order even when the task was explicitly non-communicative.

In summary, a difference in people's preferred word order for encoding or communicating meanings of reversible vs. non-reversible events would suggest that word orders are shaped by noisy-channel pressures. Here, we show exactly this pattern of performance, with gestured word orders being dependent on the semantic reversibility of the event, across three languages: an SVO language (English), and two SOV languages (Japanese and Korean).

Experiments

Participants first verbally described, and then gestured, events that involved one, two or three participants. Results are summarized in Figure 1, which shows the proportion of subjects who gestured and uttered O before V, in each experiment. Experiments with the same number (1Eng/1Jpn/1Kor; 2Jpn/2Kor) have the same design across languages; the suffix indicates the language.

[Figure 1.]

For Experiments 1Eng/1Jpn/1Kor and 2Jpn/2Kor, we consider three (not mutually exclusive) factors that might affect the order of participants' gestures: (1) an initial bias in favor of SOV order (Goldin-Meadow et al., 2008); (2) an initial bias in favor of the word order of the participant's native language; and (3) communicative or memory pressures in the form of a noisy-channel model. In an SVO language – as in Experiment 1Eng – both factors (2) and (3) predict a shift to SVO order for reversible events from the baseline SOV order, but for different reasons. In an SOV language – as in Experiments 1Jpn / 1Kor / 2Jpn / 2Kor – only factor (3) predicts a shift to SVO order for reversible events. In Experiment 3Eng, we investigate an alternative to the noisy-channel hypothesis based on minimizing syntactic dependency distances.

Methods

38 native English speakers (Experiment 1Eng, n=25; Experiment 3Eng, n=13), 23 native Japanese speakers (Experiment 1Jpn, n=11; Experiment 2Jpn, n=12) and 24 native Korean speakers (Experiment 1Kor, n=12; Experiment 2Kor, n=12) participated for payment. Participants were excluded for: knowing sign language (n=1) or failing to

follow instructions (n=3), leaving 34 (21 male) English, 23 (6 male) Japanese, and 24 (17 male) Korean participants.

Participants watched brief silent animations of intransitive and transitive events. First, participants verbally described each vignette. Then participants watched the vignettes again, in the same order, and gestured the event meanings (Figures 2-3). Participants were informed that their gestures would be filmed, and they were asked to use hand-gestures only. Participants readily completed the gesture task with minimal instruction. All responses were video-recorded and coded offline by two independent coders.

[Figures 2 and 3.]

Verbal and gesture responses to each vignette were coded for the relative position of the agent, action and patient. Trials where participants did not mention the patient, or mentioned the patient or the action in more than one position were omitted from the analyses (Expt 1Eng: 9.7% of the trials; Expt 1Jpn: 5.1%; Expt 1Kor: 5.2%; Expt 2Jpn: 6.3%; Expt 2Kor: 7.8%; Expt 3Eng: 3.4%). Inter-coder agreement about the order of the agent, action and patient was 95% across experiments. If the coders disagreed, the primary experimenter's judgment for each experiment was used: Kimberly Brink for Experiment 1Eng/3Eng and Eunice Lim for Experiments 1Jpn/1Kor/2Jpn/2Kor.

Experiment 1Eng: English (SVO) participants

In Experiment 1Eng, we varied the patients of transitive events between human and inanimate entities, so that the sentences were either semantically non-reversible or reversible (Figure 2). If gesture production is sensitive to reversibility of the event, then more SVO word orders should be produced for events where both event participants are human, and thus are equally likely to be the agent or patient. Participants saw 8 transitive events with inanimate patients (e.g., "girl kicks ball"), 8 transitive events with human patients (e.g., "girl kicks fireman"), and 8 intransitive distractor events. The same 8 actions were used for the human and inanimate patients (pushing, poking, kissing, throwing, kicking, rubbing, elbowing, and lifting).

Results

In their verbal responses participants uniformly used English word order (SVO). Replicating Goldin-Meadow et al. (2008), participants generally gestured the patient before the action when the patient was inanimate (68% of trials). However, they generally gestured the action before the patient when the patient was human (71% of trials), as predicted by the noisy-channel hypothesis. This difference (68% verb-final vs. 29% verb-final) was statistically significant in a one-tailed mixed-effects logistic regression with participant slopes and intercepts (β= 2.57, z=5.25, p < .001) (Gelman & Hill, 2007). (This test is used for all results to follow[2], except when many participants are near 0 (or 1). In such cases, logistic regression is inappropriate and the models do not converge; for these contrasts, we present Wilcoxon paired comparisons.) Although

---

[2] All theoretically relevant results that are significant with this test are also significant at p<0.05 with a one-tailed paired Wilcoxon test computed on individual subject proportions for each condition.

human patients were gestured before the action on a minority of trials, this proportion was still significantly higher than in the verbal condition (29% vs. 0%; $p<0.001$).

Experiments 1Jpn/Kor and 2Jpn/Kor: Japanese and Korean (SOV) participants

The results of Experiment 1Eng in English can be explained by the combination of the SOV default and the native-language word order prior, without invoking the noisy-channel hypothesis. In particular, participants may shift to their native language when the materials are reversible, potentially in response to increased ambiguity. We therefore used the same method and materials as in Experiment 1Eng with participants of two SOV languages: Japanese (Experiment 1Jpn) and Korean (Experiment 1Kor). To the extent that the English shift from SOV from SVO is due to communicative / memory pressures as predicted by the noisy-channel hypothesis, then Japanese and Korean speakers should also shift order to SVO for reversible items, although their native language is SOV.

For Experiments 2Jpn and 2Kor we used more complex materials: the events from Experiment 1 embedded in a "thought" or "utterance" bubble (e.g., Figure 3 conveys that the old woman says that the fireman kicks the girl; see Langus & Nespor, 2010, for a similar design but without manipulating the reversibility of the event in the embedded clause). These more complex constructions provide an even stronger test for the native language word order hypothesis vs. the noisy-channel hypothesis. If participants simply use their native language word order when materials get ambiguous or otherwise more complex, then Japanese and Korean speakers should gesture these events with the SOV order: S_1 [$S_2O_2V_2$] V_1 (e.g., "woman [fireman girl kicks] says"). However, for reversible events, in which all three event participants are human, this word order creates maximum potential confusion according to the noisy-channel hypothesis. So, if participants aim to create the most robust-to-noise representation of an event, then Japanese and Korean speakers may gesture these materials using the SVO order: S_1V_1 [$S_2V_2O_2$] ("woman says [fireman kicks girl]").

Results

In both Experiments 1Jpn and 1Kor, participants always verbalized the patient before the action (100%) and behaved similarly in their gestures: they gestured the patient before the action independent of the animacy of the patient (Japanese: 99% for inanimate patients, 95% for human patients, Wilcoxon $p=0.25$; Korean: 99% for inanimate patients, 97% for human patients, Wilcoxon $p=1.0$). These results are consistent with a role for the native language prior.

Critically, in Experiment 2Jpn and 2Kor, both Japanese and Korean participants gestured the top-level verb in second position, between the top-level subject and embedded subject in 99% of the trials, as compared to 0% (Japanese) and 23% (Korean) in verbal descriptions (Wilcoxon $p<0.005$ for each), where the top-level verb was produced sentence-finally.[3]

In the embedded clause, human patients were gestured before the action in only 66% (Japanese) and 57% (Korean) of vignettes, compared to 85% (Japanese) and 86% (Korean) of vignettes with an inanimate patient (Japanese: $\beta= 1.56$, $z=1.74$, $p < .05$;

[3] The 23% of trials that were produced with the top-level verb in second position by Korean participants were actually produced as two sentences, as evidenced by the presence of the verbal suffix "-da" – a formal politeness pragmatic mood marker for top-level clauses – following each clause (e.g., "[Boy says]-da. [Girl heart pokes]-da.").

Korean: $\beta = 3.01$, $z=2.88$, $p < .005$), as predicted by the noisy-channel hypothesis. That is, Japanese and Korean participants gestured SVO order for materials with human patients 34% and 43% of the time, respectively. Each of these proportions was reliably different from the verbal condition where human patients were produced before the action on all trials in both languages. In summary then, these results are predicted by the noisy-channel hypothesis, but not by a model containing only SOV and native language priors.

Experiment 3Eng: Minimizing syntactic dependency distances?

Although the results of Experiments 1 and 2 are consistent with a noisy-channel approach to representational robustness, they are also potentially consistent with an alternative explanation: minimizing syntactic dependency distances. In particular, the memory demands of a sentence may be sensitive to the distances in terms of the number of words between a syntactic head (e.g., a verb) and its dependents (e.g., its subject/object), such that structures and languages with shorter head-to-dependent distances are easier to process, both in production and comprehension (e.g., Gibson, 1998; Hawkins, 2004; Temperley, 2007; Tily, 2010). The dependency-distance hypothesis provides an explanation for another crosslinguistic generalization: If in a language verbs precede (rather than follow) their objects – as in SVO – then prepositions generally precede their argument noun phrases, and complementizers (embedded clause markers) precede their embedded clauses (Greenberg, 1963). Thus dependency distances might also underlie a shift from SOV to SVO word order given that SVO order allows shorter dependency distances across many constructions.

To test the dependency-distance hypothesis, we varied the complexity of the descriptions of the patients of ditransitive verbs by including 0-3 salient features. Animations consisted of a boy and a girl interacting with a set of objects (a circle, a star, or a heart). Objects had 0-3 features: surface (spotted/striped), container (in a box/pail), and/or headwear (wearing a top hat/a witch's hat). 12/36 vignettes involved a “giving” event (e.g., “the girl gives the boy a ball”). 12/36 vignettes involved a “putting” event (e.g., “the girl puts a star on the table”). The remaining vignettes involved intransitive events that were the same as in Experiments 1-2. If participants are sensitive to the linear distance between the agent, the patient, and the verb, then we expect a higher rate of SVO gesture order for longer patient descriptions, because this order minimizes the dependency distances. The noisy-channel hypothesis predicts no such shift to SVO order, because the patient is not a possible agent of the verb, and because adding modifiers to the patient does not affect the recoverability of the meaning (i.e., who is doing what to whom).

Results

Participants gestured the patient before the action for 88% of ditransitive events, compared to 8% of spoken descriptions (Wilcoxon $p<.005$). Furthermore, the number of features on the patients (i.e., the length of the patients' descriptions, which were on average approximately the same as the number of features in the target objects: 0, 0.89, 1.89, 2.68 for items with 0, 1, 2, 3 features respectively) did not affect the order of the gestures in a logistic mixed-effects regression with subject slopes and intercepts, with number of features treated as a continuous ($\beta=0.006$, $t=0.3$, $p>0.38$) or a categorical ($\beta <$

0.07, $t < 0.89$, $p>0.27$ for all levels) predictor. Even when the gestures became very long and unwieldy, participants continued to gesture the patient before the action, consistent with the noisy-channel hypothesis, but inconsistent with the dependency-distance hypothesis. Participants made more errors on complex patients, omitting and mislabeling features more often (χ^2 (3)=14.81; $p<0.002$), but the relative positions of the patient and the action were not affected (χ^2(3)=2.40; $p=0.49$)

Discussion

We have proposed and evaluated a novel account for the prevalence of SOV and SVO orders, and the OV/VO crosslinguistic variation, within the framework of Shannon's theory of communication. According to this account, speakers have a default SOV word order preference, but their choice of word order is affected by the desire to maximize meaning recoverability in the face of possible noise.

We replicated a strong SOV preference in gesture production of English-speaking participants when the subject (agent) was human and the object (patient) was an inanimate object (Goldin-Meadow et al., 2008). We further extended these results by demonstrating a similarly strong SOV preference even when the inanimate patient has up to three additional features to be gestured (Experiment 3Eng). Consistent with the claims of Goldin-Meadow and colleagues, these results suggest that SOV is the preferred word order in human communication.

Critically, our results also showed that when both the agent and patient are human, the SOV order preference disappears, with participants being now more likely to use SVO word order. Although in simple materials, speakers of SOV languages (Japanese, Experiment 1Jpn; Korean, Experiment 1Kor) tended to gesture SOV order (consistent with the native language bias), in Experiments 2Jpn (Japanese) and 2Kor (Korean), participants' gestures were inconsistent with the native language bias. First, participants reliably produced the top-level verb in second position, thus separating the top-level subject (in initial position) and the embedded subject (in third position). This word order contrasts with SOV order, under which the top-level verb appears in the final position, following the embedded clause (see Langus & Nespor, 2010, for similar results in embedded materials in Turkish, another SOV language). Second, participants tended to shift to SVO order for the embedded clause depending on the reversibility of the embedded clause, as predicted by the noisy-channel hypothesis, but not by the native language bias. We propose that the shift to SVO order for semantically reversible materials occurs in order to maximize meaning recoverability, as predicted by a model of language that includes a noisy-channel communicative component (see also Hall, Ferreira & Mayberry, 2010, for similar results from English, and Meir et al., 2010, for similar results from Hebrew, another SVO language).

In addition to explaining gesture-production data, the noisy-channel hypothesis can explain four crosslinguistic typological patterns:

(1) Information sources other than word order should mitigate the confusability of the subject and object in SOV order. Case-marking is one way to mark syntactic/semantic roles. If a linguistic community invents case-marking then the noisy-channel hypothesis predicts that the default SOV order will be retained. If, however, the community

doesn't invent case-marking (or agreement, or some other way of conveying semantic roles), then the noisy-channel hypothesis predicts that they will shift to SVO order in order to communicate optimally. This proposal thus predicts that SOV languages should tend to be case-marked, while SVO languages should tend not to be case-marked. Indeed, in a description of 502 languages from across the world, 181 of the 253 SOV languages (72%) are case-marked, whereas only 26 of the 190 SVO languages (14%) are case-marked (Greenberg, 1963; Venneman, 1973; Dryer, 2002; Croft, 2002).

We can further evaluate the hypothesis that SOV word order should be case-marked in our gesture experiments if we can find a plausible gestural cue that may serve a purpose similar to case-marking. One such plausible cue is *location in space*: many gesturers would sometimes use one hand to gesture one event participant, and the other hand for the second participant in a transitive event, or they would use different locations in space for different event participants, such that one spatial cue indicated the agent, and the other indicated the patient of the action. We can then evaluate the gesture order according to whether spatial cues were used to disambiguate semantic roles or not. In a post-hoc analysis (see Table 1), we indeed observed an effect of spatial “case-marking”. For the critical reversible materials in English Experiment 1Eng, of the spatially-marked trials, 23/36 = 64% were

SOV (36% SVO); only 15/109 = 14% of non-spatially-marked trials were SOV (86% SVO). Similar results obtained in Japanese Experiment 2Jpn (of the spatially-marked trials, 40/51 = 78% were SOV vs. only 17/35 = 49% SOV for the non-spatially-marked trials), and in Korean Experiment 2Kor (of the spatially-marked trials, 18/28 = 64% were SOV vs. 32/60 = 53% SOV for the non-spatially-marked trials).

(2) Case-marking can be animacy-dependent. If case-marking resolves the communicative ambiguity that arises for reversible events, then it should be asymmetric: animate direct objects should be more likely to be case-marked than inanimate objects. Indeed, approximately 300 languages exhibit Differential Object Marking (DOM) (Aissen, 2003) in which only animate direct objects are case-marked.

(3) Word order can be animacy-dependent. In particular, among languages with relatively free word order (allowing both SOV and SVO word orders) many demonstrate "word order freezing": in reversible constructions, when case does not disambiguate semantic roles, SVO word order is preferred (e.g., Russian (Jakobson, 1936; Bouma, 2011) and Kata Kolok, a sign language in northern Bali, Indonesia (Marsaja, 2008; Meir et al., 2010)).

(4) Languages that are not SVO can have more word order flexibility. According to the noisy-channel hypothesis, a language that is not SVO (e.g., SOV or VSO) must contain mechanisms other than word order to unambiguously convey meanings of reversible sentences. Consequently, these languages do not need to use word order to disambiguate, and so can allow variability in order. Therefore, fixed word order should primarily be found in SVO languages, and non-SVO languages should generally have less rigid word order. According to Dryer (personal communication) this generalization is correct.

To conclude: postulating sophisticated innate machinery (e.g., Universal Grammar; Chomsky, 1986) may not be necessary to explain word order variation across languages. Many aspects of crosslinguistic word order variation can be accounted for by communicative / memory pressures, which also explain other properties of human languages, including the composition of the sound inventories (Hockett, 1955; Lindblom & Maddieson, 1988) and the lexica (Zipf, 1949; Piantadosi, Tily & Gibson, 2011).

References

Aissen, J. (2003). Differential object marking: Iconicity vs. economy. Natural Language & Linguistic Theory, 21 (3), 435-483.

Aylett, M., & Turk, A. (2004). The smooth signal redundancy hypothesis: A functional explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech. Language and Speech, 47 (1), 31.

Baker, M. (2001). The atoms of language. Basic books.

Bouma, G. (2011). Production and comprehension in context: The case of word order freezing. In A. Benz & J. Mattausch (Eds.), Bidirectional optimality theory. Amsterdam, The Netherlands: John Benjamins.

Brady, T.F, Konkle, T., & Alvarez, G.A. (2009). Compression in visual working memory: Using statistical regularities to form more efficient memory representations. Journal of Experimental Psychology: General, 138(4), 487-502.

Chomsky, N. (1986). Knowledge of language: Its nature, origin, and use. Praeger Publishers.

Croft, W. (2002). Typology and universals. Cambridge Univ Press.

Dryer, M.S. (2002). Case distinctions, rich verb agreement, and word order type (comments on Hawkins paper). Theoretical Linguistics, 28 (2), 151-158.

Dryer, M. S. (2005). The order of subject, object and verb. In M. Haspelmath, M. S. Dryer, D. Gil, & B. Comrie (Eds.), The world atlas of language structures (pp. 330–333). Oxford University Press: Oxford.

Gell-Mann, M., & Ruhlen, M. (2011). The origin and evolution of word order. Proceedings of the National Academy of Sciences, 108 (42), 17290-17295.

Gelman, A., & Hill, J. (2007). Data analysis using regression and multilevel/hierarchical models. Cambridge University Press New York, New York, USA.

Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. Cognition, 68 (1), 1-76.

Gibson, E., & Bergen, L. (2012). The rational integration of noise and prior semantic expectation: Evidence for a noisy-channel model of sentence interpretation. Poster presented at the 25th annual CUNY sentence processing conference. New York, NY, March, 2012.

Givón, T, (1979). On Understanding Grammar. Academic Press, New York.

Goldin-Meadow, S., So, W., Ozyurek, 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.

Greenberg, J. (1963). Some universals of grammar with particular reference to the order of meaningful elements. In Greenberg, J. (ed.) *Universals of Language*, pp. 73-113.

Hall, M., Ferreira, V.S. & Mayberry, R.I. (2010). WOMAN BOX PUSH but not WOMAN BOY PUSH: Reversibility changes argument order in pantomime. Poster presented at the Psychonomics Society.

Haspelmath, M. (1999). Why is grammaticalization irreversible? *Linguistics*, 37 (6), 1043-1068.

Hawkins, J. (2004). Efficiency and complexity in grammars. Oxford University Press, USA.

Hockett, C. (1955). *A Manual of Phonology*. Waverly Press, Baltimore.

Hockett, C. (1960). The origin of language. *Scientific American*, 203 (3), 88-96.

Jackendoff, R. (1972). *Semantic interpretation in generative grammar*. Cambridge, MA: MIT Press.

Jaeger, F. (2010). Redundancy and reduction: Speakers manage syntactic information density. *Cognitive Psychology*, 61 (1), 23-62.

Jakobson, R. (1936). Contribution to the general theory of case: general meanings of the Russian cases [translation of beitrag zur allgemeinen kasuslehre: Gesamtbedeutung der russischen kasus.] (No. 6).

Kuperman, V., Stadhagen-Gonzales, H. and Brysbaert, M. (2012). Age-of-acquisition ratings for 30 thousand English words. *Behavior Research Methods*.

Langus, A., & Nespor, M. (2010). Cognitive systems struggling for word order. *Cognitive Psychology*, 60 (4), 291-318.

Levy, R. (2008). A noisy-channel model of rational human sentence comprehension under uncertain input. In Proceedings of the conference on empirical methods in natural language processing (pp. 234-243).

Levy, R., Bicknell, K., Slattery, T., & Rayner, K. (2009). Eye movement evidence that readers maintain and act on uncertainty about past linguistic input. *Proceedings of the National Academy of Sciences*, 106 (50), 21086-21090.

Lindblom, B. and Maddieson, I. (1988). Phonetic universals in consonant systems. In Li, C. and Hyman, L., editors, Language, Speech and Mind, pages 62–78. Routledge, London.

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

Marsaja, I. G. (2008). Desa kolok - a deaf village and its sign language in Bali, Indonesia. Nijmegen: Ishara Press.

Meir, I., Lifshitz, A., Ilkbasaran, D., & Padden, C. (2010). The interaction of animacy and word order in human languages: A study of strategies in a novel communication task. In Proceedings of the eighth evolution of language conference (pp. 455-456).

Meir, I., Sandler, W., Padden, C., & Aronoff, M. (2010). Emerging sign languages. In M. Marschark & P. Spencer (Eds.), Oxford handbook of deaf studies, language, and education (Vol. 2). New York: Oxford University Press.

Newcombe, R. (1998). Interval estimation for the difference between independent proportions: comparison of eleven methods. Statistics in Medicine, 17 (8), 873-890.

Newmeyer, F. (2000a). Language form and language function. MIT Press, Cambridge, MA.

Newmeyer, F. (2000b). On the reconstruction of 'proto-world' word order. In Knight, C., Studdert-Kennedy, M. & Hurford, J. *The Evolutionary Emergence of Language: Social Function and the Origins of Linguistic Form.* pp. 372-388. Cambridge University Press.

Paul, H. (1880). Prinzipien der sprachgeschichte [Principles of the history of language]. Tuebingen, Germany: Max Niemeyer.

Piantadosi, S., Tily, H., & Gibson, E. (2011). Word lengths are optimized for efficient communication. *Proceedings of the National Academy of Sciences*, 108 (9), 3526-3529.

Piantadosi, S., Tily, H., & Gibson, E. (2012). The communicative function of ambiguity in language. *Cognition*, *122*, 280-291.

Pinker, S., & Bloom, P. (1990). Natural language and natural selection. Behavioral and Brain Sciences, 13, 707-784.

Sandler, W., Meir, I., Padden, C., & Aronoff, M. (2005). The emergence of grammar: Systematic structure in a new language. *Proceedings of the National Academy of Sciences of the United States of America, 102 (7),* 2661-2665.

Schouwstra, M. van Leeuwen, A., Marien, N., Smit, M. de Swart, H. (2011). Semantic structure in improvised communication. Proceedings of the 33rd Annual Meeting of the Cognitive Science Society (CogSci11), 1497-1502.

Senghas, A., Coppola, M., Newport, E., & Supalla, T. (1997). Argument structure in Nicaraguan Sign Language: The emergence of grammatical devices. In E. Hughes & A. Greenhill (Eds.), Proceedings of the 21st annual Boston university Conference on Language Development (Vol. 2, pp. 550-561). Cascadilla Press: Somerville.

Smith, P. (1969). Coding strategies in language. Information and Control , 14 (1), 72-97.

Temperley, D. (2007). Minimization of dependency length in written English. Cognition, 105 (2), 300-333.

Tily, H. (2010). The role of processing complexity in word order variation and change. Stanford University.

Vennemann, T. (1973). Explanation in syntax. In John Kimball (ed.), Syntax and semantics (vol. 2), 1-50. New York: Academic Press.

Zipf, G.K. (1949). *Human Behavior and the Principle of Least Effort*. New York: Addison-Wesley.

Figure Captions

Figure 1: Summary of results across Experiments 1-3. For all experiments, we show the proportion of patient-before-action productions (OV) in dark grey vs. action-before-patient productions in white (VO). For Experiment 2, the lower two panels show the proportion of productions in which the top-level verb was produced in second position (in light grey) vs. the proportion of productions in which the top-level verb was produced in final position (in white). The critical reversible vs. non-reversible gesture patterns are circled for Experiment 1Eng, 2Jpn, 2Kor.

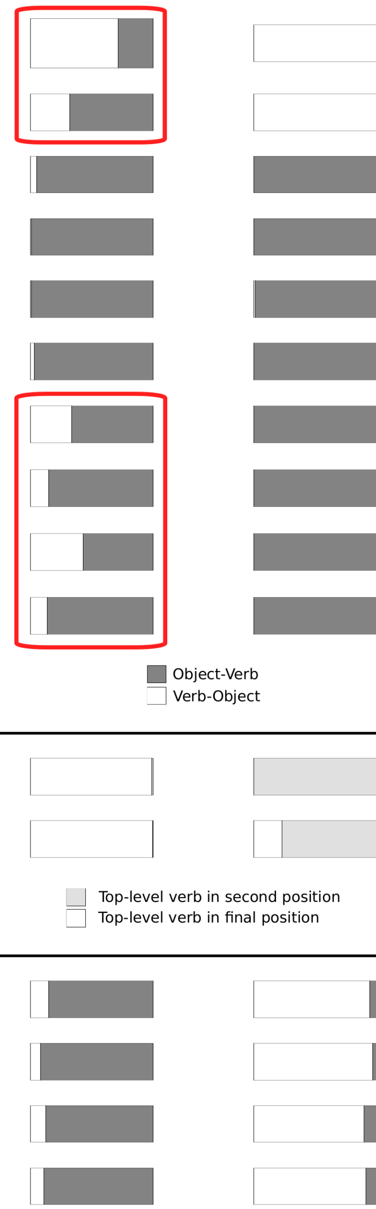
Figure 2: Sample trials from Experiment 1: The top panel shows an inanimate patient condition (non-reversible event), and the bottom panel shows an animate patient condition (reversible event).

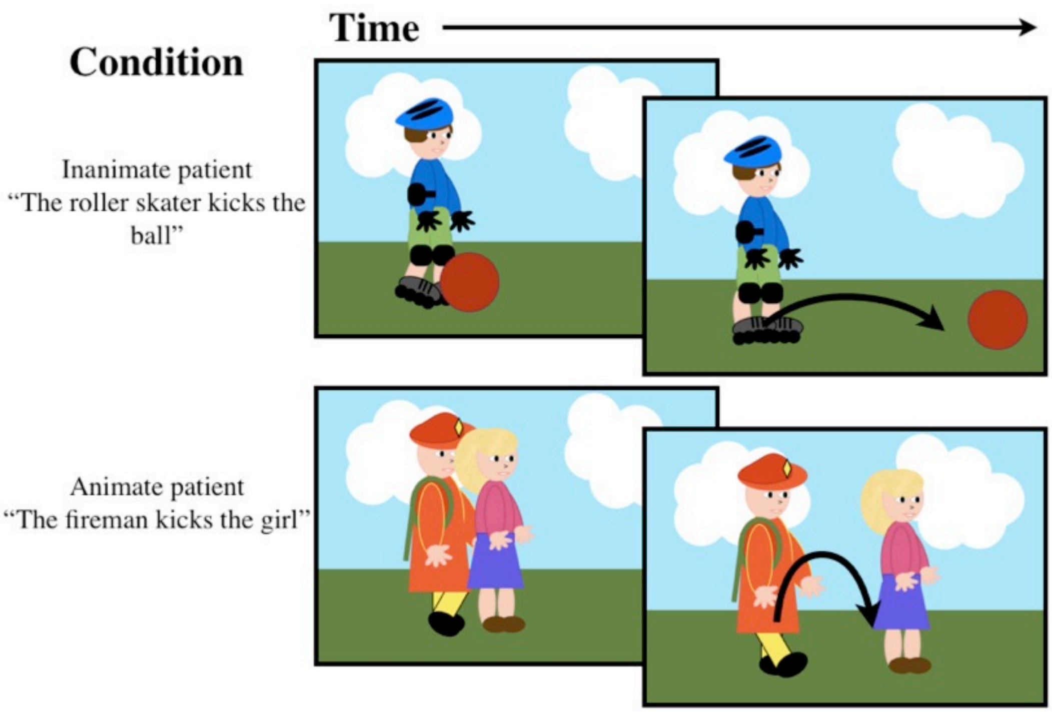
Figure 3: A sample trial from Experiment 2.

Experiment	1Eng		2Jpn		2Kor	
Word order	**VO**	**OV**	**VO**	**OV**	**VO**	**OV**
			No spatial case-markers			
Non-reversible	38	26	6	26	5	31
Reversible	94	15	18	17	28	32
			+ spatial case-markers			
Non-reversible	18	93	7	48	7	43
Reversible	13	23	11	40	10	18

Table 1: The distribution of trials in Experiments 1Eng, 2Jpn and 2Kor when spatial case-markers were provided to disambiguate the semantic roles of the participants in the events. Critically, note that for the reversible materials, people were relatively more likely to gesture VO word order when they did not provide spatial case-markers.

Experiment	Language	Manipulation	Gesture	Verbal
1Eng	English	Reversible		
		Non-reversible		
1Jpn	Japanese	Reversible		
		Non-reversible		
1Kor	Korean	Reversible		
		Non-reversible		
2Jpn	Japanese	Embedded Reversible		
		Embedded Non-reversible		
2Kor	Korean	Embedded Reversible		
		Embedded Non-reversible		
2Jpn	Japanese	Top-level		
2Kor	Korean	Top-level		
3Eng	English	1 Feature		
		2 Features		
		3 Features		
		4 Features		




Time
Condition
Inanimate patient
“The roller skater kicks the ball”
Animate patient
“The fireman kicks the girl”

Condition

Animate patient
in embedded clause
"The old woman says that
the fireman kicks the girl"