Informational sensitivity and communicative order

Andrew Shibata

June, 2017

Language and Cognition Lab, UC Berkeley

Advisors: Terry Regier, Alexandra Carstensen

Second Reader: Susanne Gahl

Abstract

In 2008, Goldin-Meadow et al. claimed that SOV is the ‘Natural Order of Events’ with regards to gestural communication order. Our study addresses potential confounds concerning task-based and corpus-based informativeness. In our study’s experiments, English speakers played a communication game to assess the influence of communicative pressure, English Language Statistics, and Game Statistics on gestural communication order. Our results indicate that people attend to relative informativeness given communicative pressure, but that they do not use relative informativeness to determine communicative order.

Introduction

Word and gesture order

A basic clause contains a Subject (S), Object (O), and Verb (V). For languages with a dominant word order, the two most prevalent word orders across languages are SOV [47.6%] and SVO [41.0%] (Dryer, 2005). The remaining possible orders combined (OSV, OVS, VSO, VOS) are only attested in 10.4% of such languages. This distribution suggests that humans tend to prioritize Subjects in their utterances. How languages subsequently developed Object-Verb or Verb-Object order is less clear.

Besides spoken language, humans demonstrate co-speech gesture which either reflects or enhances the content conveyed through speech. Two stances that many linguists take regarding the role of gesture in communicative discourse are (1) communicators gesture to help facilitate utterance production or (2) communicators gesture with the explicit intent of conveying information to their interlocutors (Alibali, Heath, and Meyers, 2001). Our study focuses on (2). We hypothesize that language producers order communication based on relative informativeness to assist the interlocutor in recovering the meaning from the signal. While gesture mirrors the content presented in spoken language, Goldin-Meadow et al. (2008) demonstrated that gesture order does not covary with spoken language constituent order.

Goldin-Meadow, So, Özyürek, and Mylander (2008) studied spontaneous, co-speech gesture production in English (SVO), Spanish (SVO), Turkish (SOV), and Mandarin (SVO) speakers. Goldin-Meadow et al. used two tasks to study gesture order: a gesture task in which subjects described vignettes only through gesture, and a transparency task in which subjects reconstructed vignettes by stacking transparent pictures on a peg. They found that people had a bias for SOV gesture order regardless of the speaker's native language word order. Their most striking claim is that SOV is the "Natural Order of Events" that humans use to temporally frame relationships between agents and objects.

Challenges to the natural order view

In response to the claim that event structure follows a "natural" SOV order, researchers demonstrated that gesturers deviate from SOV order when describing events with certain characteristics. Gesture order preference depends on the semantic properties of the verb. Gesturers prefer SVO order when gesturing events with intensional verbs (Schouwstra & de Swart, 2014) and interpret ambiguous SVO events as intensional rather than extensional (Schouwstra, van Leeuwen, Marien, Smit, & de Swart; 2011). Gesture order favors SVO order when describing semantically reversible events. Gibson et al. (2013) argue that SVO order is useful for temporally separating reversible agents and patients in noisy environments. Hall, Mayberry, & Ferreira (2013) make a similar argument, but claim that the semantic confusion comes from patient-verb adjacency. Langus & Nespor (2010) try to reconcile findings supporting both SOV and SVO preferences by asserting that the cognitive systems for computational grammar and improvised communication prefer different orders.

In addition to the empirical challenge to the natural order view, there is a theoretical challenge from an efficient communication perspective. Naturalness and efficient communication are not mutually exclusive, but they are confounded through imbalanced informativeness in the Goldin-Meadow et al. (2008) stimuli. Their claim that this bias is independent of language structure serves as a motivation to study gesture order preference through the lens of information theory:

> "…the ordering we use when representing events in a nonverbal format is not highly susceptible to language's influence. Rather, there appears to be a natural order that humans (regardless of the language they speak) use when asked to represent events nonverbally" (p. 9167).

Efficient communication is a broad approach to cognitive linguistics asserting that languages are structured to be maximally informative as well as simple (Regier, Kemp, and Kay, 2015). Efficient communication is supported by studies demonstrating that languages become more progressively efficient as they are transmitted to later generations (Carstensen et al., 2015) and that words maximize the amount of information conveyed per unit of length (Piantadosi, Tily, and Gibson, 2011). This study explores the domain of communicative word order through the perspective of information theory. Efficient communication would predict that speakers order utterances using a simple function of informativeness. Communicating more informative tokens earlier can help the interlocutor predict which information will follow in an utterance.

Many studies have used information theory to study communication. Zipf (1935) demonstrated the inverse relationship between word frequency and rank of frequency. Jaeger & Tily (2011) established that word order considers simplicity and word expectations. Gibson et al.

(2013) used Shannon’s model of communicative uncertainty, or entropy (1948), to argue that SVO gesture order helps interlocutors recover information when noise is present.

Maurits (2012) most directly studied the relationship between word order and information theory. Maurits examined the Uniform Information Density (UID) Hypothesis (p. 132) which claims that language producers try to produce maximally functional language by distributing uncertainty (Shannon’s entropy) uniformly throughout communication. This hypothesis suggests that language producers aim to structure communication so that the interlocutor receives information at a uniform, non-surprising rate. Maurits analyzed several English corpora and conducted an experiment on the perceived likelihood of transitive events to conclude that Objects are more informative than Verbs in English (for Verb-Informative versus Object-Informative examples, see **Appendix I**).

While Maurits asserts that information is uniformly distributed throughout communication, we hypothesize that language producers order words from most informative to least informative. Under this hypothesis, language producers front important information in communicative order to provide interlocutors with sufficient discursive context early in the utterance. One communicative domain that supports this hypothesis is the existence of “no gap, no overlap” in communicative turn-taking (Sacks et al., 1974; Roberts et al., 2015). One interaction within conversations involves one interlocutor beginning their turn immediately after, or even before, the other finishes their turn. The interlocutor does not process the concluding portions of the initial signal. A language producer may intend to front their utterance with the most informative constituents since the interlocutor is more likely to be attending to that portion of the utterance. In summary, we hypothesize that word order is determined as a monotonic function of informativeness – from most informative constituents to least informative constituents.

The naturalness of OV order from the Goldin-Meadow et al.'s (2008) stimuli is confounded by informativeness in several ways:

(1) Their task-stimulus frequencies were imbalanced between Objects and Verbs. They repeated Verbs ('gives', 'carries') between stimulus sentences, but did not repeat any Object Nouns. Since Verbs appeared multiple times within the stimulus pairs, Object Nouns provided more contrastive information identifying which vignette the subjects were communicating.

(2) The conditional probabilities for their Object-Verb stimulus pairs were biased towards Noun-informative pairs. We evaluated Goldin-Meadow et al.'s stimulus sentences for transitive sentences using conditional probability calculations from COCA (see **Appendix II** for further information). Of the 16 stimulus pairs, 13 demonstrated the relationship *P(Verb | Object)* > *P(Object | Verb)*, indicating that the Objects provided more information about the Verbs. For the full table of calculations, see **Appendix III**.

(3) The conditional probabilities for Verbs and Objects in English transitive sentences are generally biased towards Object-informativeness. Maurits' (2015) corpus analysis demonstrated that "…there are more objects in the world than there are actions" (p. 174). He found that, generally in English, *P(Verb | Subject, Object)* > *P(Object | Subject, Verb)*. Since there are more Objects that a language producer can refer to, knowing the target Object provides more information about which Verb to expect than the same target Verb provides about the collocate Object.

(4) The conditional probabilities for Verbs and Objects in English perception of event likelihood are biased towards Object-informativeness. Maurits (2015) experimentally

demonstrated that English speakers perceive Verb-Object order to be more likely than Object-Verb orders (p. 161). This perceptual order bias seems to conflict with Goldin-Meadow et al.'s (2008) claim that people are biased for Object-Verb event structure.

Our Study

For the potential informativeness confounds listed above, it is only feasible to address (1) and (2) since it is not possible to change Object and Verb frequencies in the English lexicon or change English speakers' perceptions of event likelihood. We addressed (1) by counterbalancing conditions in our experiment for which our "Game Statistics" were biased towards Objects, Verbs, or neither. We addressed (2) by similarly counterbalancing conditions in our experiment for which our "English Language Statistics" were biased towards Objects or Verbs by using stimulus pairs for which the COCA conditional probabilities were biased towards more informative Objects or Verbs.

In our study, we placed subjects in a communication game to establish communicative pressure. Since their goal was to help their partner select the correct sentence card, they were led to consider which word was most helpful for their game partner. We varied the relative frequencies of Nouns and Verbs in our stimulus sentences to manipulate "Game Statistics" and selected Object-Verb pairings based on relative COCA conditional probabilities to manipulate "English Language Statistics".

Our study asks the question "Are gesturers sensitive to informativeness?" for which we conclude yes. From this we ask if informativeness drives gesture order to which we conclude no; the informativeness confounds (1) and (2) that we addressed cannot account for the SOV bias observed in Goldin-Meadow et al. (2008).

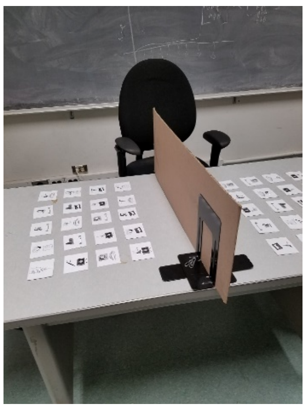
General procedure

The overarching experimental design involved 2 phases: the learning phase, and the communication game phase.

In the learning phase, subjects watched a set of videos that showed a card with an image representing a stimulus word, played audio naming the displayed word, and, after a brief delay, a hand gesture associated with the word. Subjects watched these videos twice before being tested to ensure mastery of the sign language. In this test, subjects were shown each video including the image and the word audio before the video was paused. Subjects were instructed to produce the sign and a subject was considered a master of the sign language if they produced all correct signs for 3 runs of videos in a row. Afterwards, the subjects moved from the computer seat to a different table for the communication game.

For the communication game, subjects sat on one end of a table with the game partner (confederate) on the other end. Both had an array of 'visual sentence' cards consisting of an Object-Verb word pairing and there was a large divider in the middle of the table so neither player could see the other's cards. Subjects were told that the experimenter would point to several cards and the subject was playing a game by helping the partner guess which card the experimenter was pointing to only using the sign language learned earlier. In most experiments, this was split into 2 rounds: in Round 1, the subject was only allowed to choose one word or the other to best help the partner guess the card; in Round 2, the subject gestured both words in the 'visual sentence'. This 2-round paradigm placed communicative pressure on the subject so that the subject would be looking for salient features to better help the game partner guess the correct card. The confederate recorded which signs the subject produced in Round 1 and Round 2.

The response variable for Round 1 is whether the subject responded Noun modally within the 5 (3 in Exp 2a) trials and the response variable for Round 2 is whether the subject responded Noun-Verb modally within the 5 (3 in Exp 2a) trials. In the “interesting” conditions (see **Table 1** below), the Round 1 modal response assessed whether subjects were sensitive to informativeness within the stimulus set. The Round 2 modal communicative order assessed whether subjects used informativeness to determine communicative order.



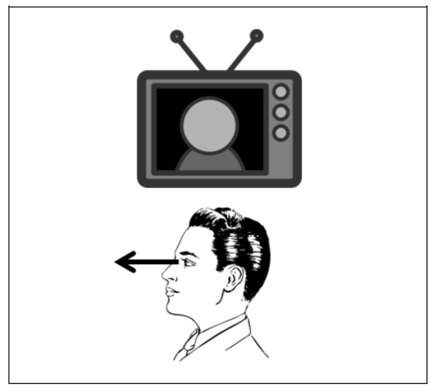


Image 1: Layout of communication game table (Experiment 2b)
Image 2: Sample visual sentence card ‘Watch TV’ for Noun Top, Verb Informative condition

Factors that we considered when designing conditions for this experiment are: ease of word-gesture memorability, order of images on stimulus cards, and informativeness direction. To avoid confounds due to specific word-gesture mappings, each experiment involved two randomized word-gesture mappings. We assigned each subject one of the two word-gesture mappings and a randomized training stimulus order. We counterbalanced image order on the cards to control for salient order. Table 1 describes the 4 major conditions in this experiment.

Table 1: Experimental Conditions

	Noun Informative	**Verb Informative**
Noun on Top of card	Noun Top, Noun Inf. (NtN)	Noun Top, Verb Inf. (NtV)
Verb on Top of card	Verb Top, Noun Inf. (VtN)	Verb Top, Verb Inf. (VtV)

(Highlighted cells are the ‘interesting’ conditions because the informative word does not match the top, or most salient, word on the card)

The informativeness of the stimuli were determined using two factors: English Language Statistics and Game Statistics. For example, in Verb Informative conditions, the stimulus word pairs were biased towards words with COCA conditional probabilities in the direction $P(O \mid V) > P(V \mid O)$ [English Language Statistics]. Additionally, in the array of visual sentences, each unique Verb appeared 2 times each while each unique Noun appeared 4 times each. If the subject selected a Verb for a given visual sentence in Round 1, the game partner would have a ½ chance of guessing the correct card on their end, but if the subject selected a Noun, the partner would only have a ¼ chance of guessing the correct card [Game Statistics] (See **Image 3** below).

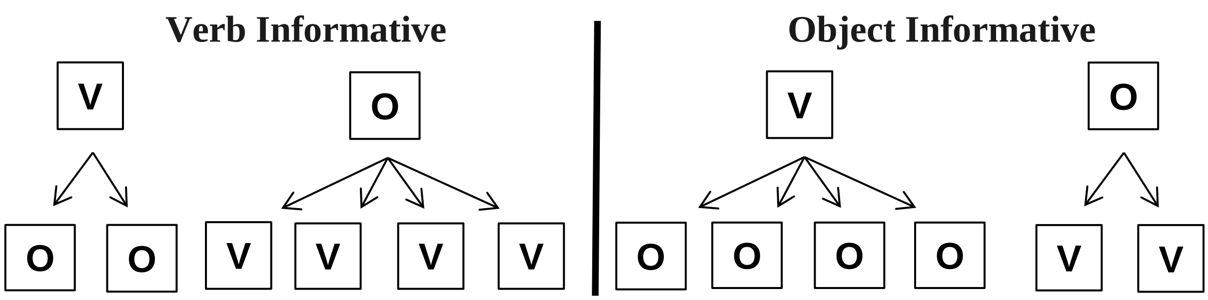


Image 3: Informativeness due to Game Statistics. Verbs and Objects appeared with imbalanced relative frequencies in the array of visual sentences.

Hypotheses

Two hypotheses for how subjects might determine which order to gesture in Round 2 are:

- **Natural Order Hypothesis (NOH):** Subjects are expected to be unconditionally biased for SOV order. 100% of subjects should gesture Object-Verb order.
- **Informative Word Hypothesis (IWH):** Subjects gesture the most informative word first and the less informative word second.

The Natural Order Hypothesis and the Informative Word Hypothesis are at odds in the NtV and VtV conditions because the NOH would expect subjects to prioritize Nouns regardless of informativeness while the IWH would expect subjects to front Verbs in communicative order.

Stimuli

(For the complete set of stimulus words, see **Appendix IV**)

We began by choosing high frequency Verbs from COCA's list of the top 5000 English words (Davies, 2008). After selecting 5 Verbs for the Verb Informative conditions and 10 Verbs for the Noun Informative conditions, we found Object Nouns that had the desired conditional probability directions [$P(V|O) < P(O|V)$ for Verb-inform, $P(V|O) > P(O|V)$ for Noun-inform]. Additionally, we calculated entropy values as defined in Shannon (1948) and chose Verbs and Nouns with comparable entropies.

The images used for each stimulus word were images that had licenses marked for reuse with modification. Each image was a black and white line drawing so that the art style was similar between stimulus images. We added an arrow to images representing Verbs to indicate that an action was occurring in the image.

The video and audio files for the training stimuli were recorded during a single session using an HTC One cell phone camera and a Blue Snowball microphone. The video and audio components were edited together using Windows Movie Maker so that each video was 5 seconds long with the hand gesture portion of the video beginning at 2.5 seconds into the video.

Each subject was assigned a training video randomization determined by performing samples in R Studio. The videos were presented in full screen mode in Windows Media Player on a Windows-based computer in the lab.

Methods

Participants

There were 188 included subjects (40 – Exp 1a, 41 – Exp 1b, 15 – Exp 2b, 47 – Exp 2a) with a mean age of 20.60 and 46 excluded subjects (29 – did not begin learning English by age 3, 9 – required greater than 12 trials to learn sign language, 6 – forgot 3 or more signs during communication game, 2 – failed to follow experimenter's instructions). All participants were UC Berkeley undergraduate students participating for research participation credit for coursework.

Experiment 1a

We tested all 4 conditions (NtN, NtV, VtN, and NtN) in Experiment 1a to establish our baseline to compare with subsequent experiments. The only difference between the general procedure described above was the inclusion of a frequency counting worksheet. After testing subject's proficiency with the sign language and before conducting the communication game, we asked subjects to count the number of times each word appeared in the set of visual sentence cards. We did this to explicitly draw attention to the imbalanced frequencies between Verbs and Nouns. We additionally randomly assigned subjects one of two randomized worksheet orders to control for potential order effects.

For Experiment 1a, we are most interested in the Round 1 results to address the question "Are gesturers sensitive to informativeness?" when all these factors are stacked to make informativeness differences salient. Requiring subjects to choose a single word in Round 1 makes the role of informativeness in determining Round 2 order unclear so we will not report those results here.

Results and discussion

For Round 1, the proportion of subjects that responded Noun-modally differed between the Verb Informative (30%) and Noun Informative (90%) conditions. This difference is significant ($\alpha = 0.05$) using a Welch two sample t-test ($t = -4.77$, $p < 0.001$). **Image 4** displays the proportion of Noun modal responses by informativeness condition.

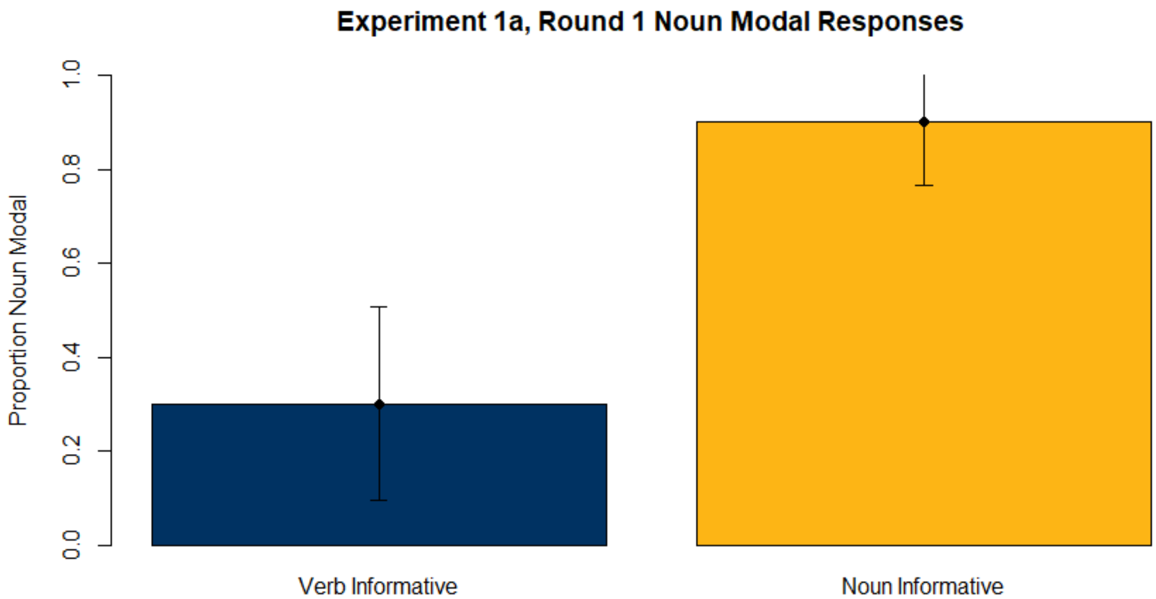


Image 4: Proportion of Noun Modal Responses in Exp. 1a, Round 1. Error bars indicate 95% confidence intervals.

The Round 1 results indicate subjects could identify the informative word in the visual sentence and used it to best help their game partner guess the stimulus sentence. Subjects select Verbs more frequently in the Verb Informative conditions when the Game Statistics and English Language Statistics are biased towards Verbs and when they explicitly count that Noun and Verb frequencies are imbalanced.

Experiment 1b

The results from Experiment 1a, Round 1 suggest that people are sensitive to word informativeness when their attention is explicitly drawn to imbalanced relative frequencies. For Experiment 1b, we repeated Experiment 1a except without providing subjects with the frequency counting worksheet. We wanted to see whether subjects were sensitive to informativeness without priming. Similarly with Experiment 1a, Round 1 addresses the question regarding sensitivity to informativeness and Round 2 results involve bias from prior response.

Since the VtV and NtN conditions do not piece apart the Natural Order Hypothesis and the Informative Word Hypothesis and since the Experiment 1a results were strongly in the Verb and Noun directions, respectively, we only explored the 'interesting' NtV and VtN conditions.

Results and discussion

There is no difference between the proportion of subjects that responded Noun-modally between Experiments 1a and 1b for the Verb Informative condition ($t = 0$; $p = 1$, Welch two sample t-test) or the Noun Informative condition ($t = 0.2325$; $p = 0.8187$, Welch two sample t-test). **Image 5** displays the Noun-modal responses for both experiments for the NtV and VtN conditions, respectively.

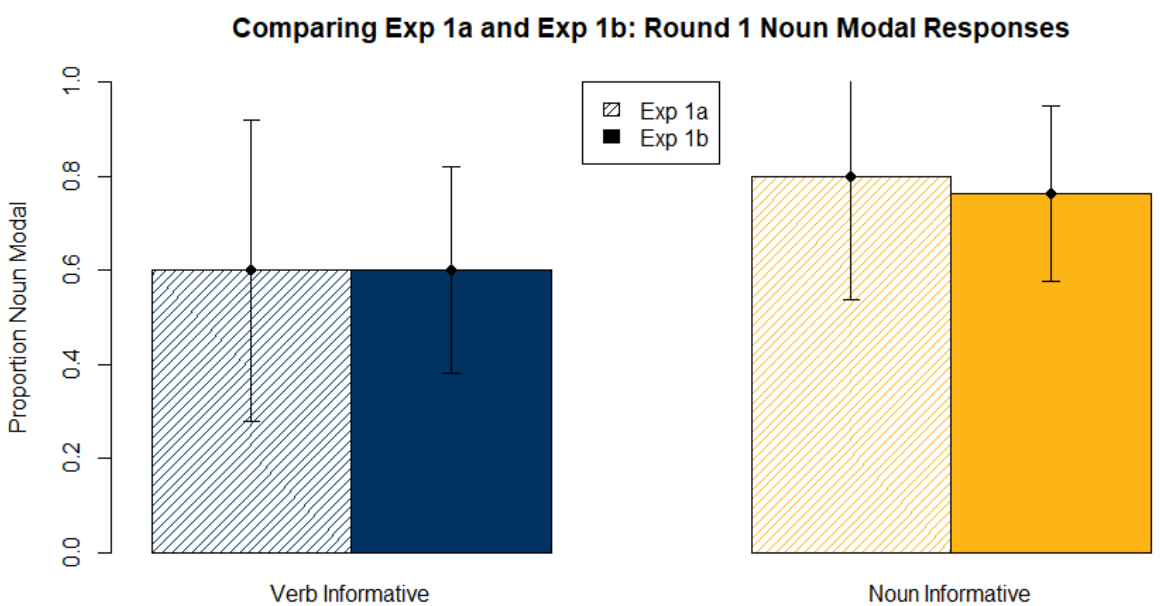


Image 5: Comparing Experiments 1a and 1b, Round 1 Noun modal responses for the NtV and VtN conditions. Error bars indicate 95% confidence intervals.

There was no difference between Round 1 responses between Experiment 1a and Experiment 1b. Since the only change was that subjects were not explicitly directed to count imbalanced Game Statistics, we conclude that drawing attention to Game Statistics-based informativeness does not affect sensitivity to informativeness.

Experiment 2a

The second set of experiments addresses the question “Can informativeness explain gesture order?”. Experiment 2a clarifies the Experiment 1 results by isolating English Language Statistics as the sole source for informativeness.

In the communication game, Noun or Verb informativeness was influenced by two factors: the English Language Statistics from the corpus analysis and the Game Statistics relative frequencies. In all the Experiment 1 designs, both English Language Statistics and Game Statistics were biased in the same direction for either Noun-informativeness or Verb-informativeness. For this experiment, we used a similar, but different set of stimulus words and did not stack all the factors in the same informative directions. We set Noun and Verb relative frequencies equal within the stimulus set, but retained the English Language Statistics biases. By setting Game Statistics equal, but keeping English Language Statistics disparate, Experiment 2a assesses whether English Language Statistics is a salient informative feature on its own.

Experiment 2a used the Round 1/Round 2 paradigm across all 4 possible conditions. For both the original Noun-informative and Verb-informative stimulus word sets, we chose 3 Nouns and 3 Verbs so that all 9 possible pairs were semantically plausible. We then chose 3 target Noun-Verb pairs that were biased towards the desired informativeness direction.

As with Experiments 1a and 1b, we are interested in the Round 1 results to investigate whether subjects are sensitive to corpus-based informativeness.

Results and analysis

Subjects did not respond differently between the Verb Informative condition (36%) and the Noun Informative condition (64%) for Experiment 2a. This difference is not significant using a Welch two sample t-test (t = -1.9246, p = 0.0607). **Image 7** compares the Round 1 responses between Experiment 1a and Experiment 2a.

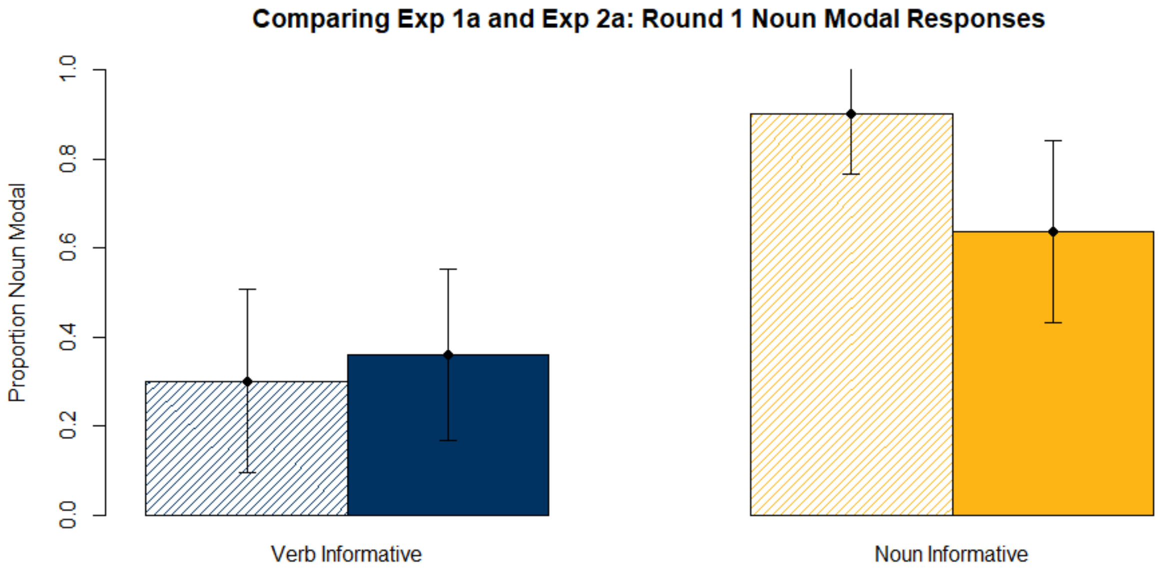


Image 7: Comparing Experiments 1a and 2a, Round 1 results. Error bars indicate 95% confidence intervals.

Qualitatively, the Experiment 2a results seem to trend towards the informative words compared to chance response, but the difference is not statistically significant. When the English Language Statistics are the only source of informativeness, subjects are not sensitive to informativeness. This suggests that subjects are sensitive to Game Statistics-based informativeness, but not English Language Statistics-based informativeness. From this we conclude that the sensitivity to informativeness we observe in Experiment 1 is being driven by the Game Statistics-based informativeness, not the English Language Statistics.

Experiment 2b

Experiment 2b returned to the same stimuli as Experiment 1, but removed the communicative pressure aspect of the communication game by removing the Round 1/Round 2 paradigm. When introducing the communication game to the subject, the experimenter simply instructed the subject to "communicate the visual sentence using the gestures that you just learned”. In this sense, subjects only provided a “Round 2” response that was either VO or OV. Without Round 1, we do not directly assess if subjects were sensitive to informativeness.

This experiment only analyzed the ‘most interesting’ condition, NtV in which the Informative Word Hypothesis and the Natural Order Hypothesis are at odds with each other.

Results and analysis

Subjects overwhelmingly used OV order (93.3%). The 95% confidence interval for this proportion (0.807, 1.06) does not capture the expected proportion assuming chance response (0.5). When comparing Experiment 2b’s results to the analogous NtV results from Experiment 1a, there is no significant difference in OV modal response proportion ($t = -0.277$; $p = 0.7849$, Welch two sample t-test). **Image 6** compares the OV modal responses for the NtV conditions in Experiments 1a and 2b.

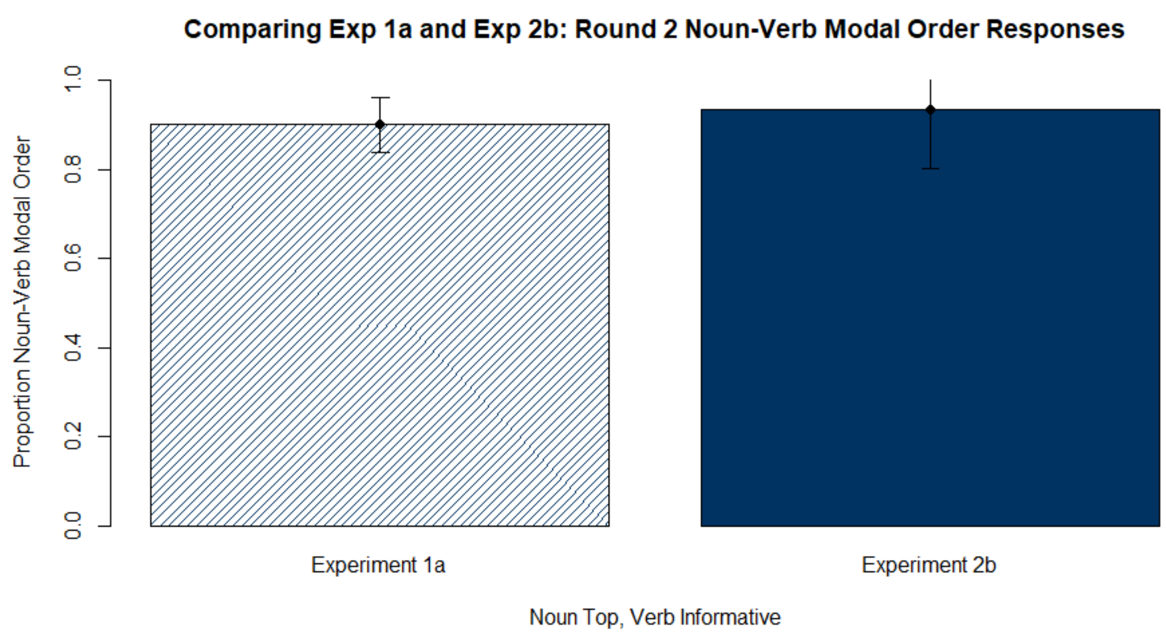


Image 6: Comparing OV modal responses for Round 2 in Experiments 1a and 2b. Error bars indicate 95% confidence intervals.

Without the communicative pressure forcing subjects to choose one word or the other, they simply gestured the most salient order: Top to Bottom, or in this case, OV order. While we did not assess whether subjects were sensitive to informativeness in this experiment, we can conclude that subjects did not order communication based on informativeness. Since informativeness does not drive communicative order, our results cannot account for the Goldin-Meadow et al. (2008) findings.

General discussion and conclusions

Neither of the hypotheses that we laid out for Round 2 responses (the Informative Word Hypothesis and the Natural Order Hypothesis) truly explained our observed Round 2 responses. Subjects did not use informativeness to determine Round 2 response order, so the IWH cannot account for the results. Subjects conditionally responded in Verb-Noun order, so the NOH cannot account for the results either. While we counterbalanced stimulus card word order, we found that subjects highly conformed to Top-to-Bottom order for their Round 2 response. Future work should involve task designs that are cognizant of visual salience as an influential factor in subjects determining communicative order. Additionally, our initial hypothesis that language producers order communication from most informative constituents to least informative constituents does not hold since our subjects did not use informativeness to order communication.

Our results suggest that people are sensitive to relative informativeness, but only when communicative pressure demands that they scrutinize the context for informative cues. While we find support for discursive informativeness sensitivity through the Game Statistics, our results do not support the claim that people are sensitive to English Language Statistics-based informativeness. We controlled for (1) imbalanced Noun and Verb relative task frequencies and (2) imbalanced stimulus collocate informativeness, but the two other confounds, (3) Object bias in general English transitive sentences and (4) Object-informative bias in the perception of events in English, may still be at play. Since we find that people are sensitive to informativeness, it is plausible that informativeness still can account for the Goldin-Meadow et al. (2008) results, but not via the confounds we investigated.

Acknowledgements

We thank Lillian Lawrence, Elizabeth Kon, and Vicki Guan for work on stimulus design as well as Kyle Eckert, James Wang, Connor Brown, and Jaguar Jung for data collection work.

References

Alibali, M. W., Heath, D. C., & Myers, H. J. (2001). Effects of visibility between speaker and listener on gesture production: Some gestures are meant to be seen. *Journal of Memory and Language*, *44*(2), 169-188.

Carstensen, A., Xu, J., Smith, C., & Regier, T. (2015). Language evolution in the lab tends toward informative communication. In *Proceedings of the 37th Annual Meeting of the Cognitive Science Society, Austin, TX, July 2015* (pp. 303-308).

Davies, M. (2008). *The corpus of contemporary American English*. BYE, Brigham Young University.

Dryer, M. S. (2005). Order of subject, object, and verb. *The world atlas of language structures*, 330-333.

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*, 0956797612463705.

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.

Jaeger, T. F., & Tily, H. (2011). On language ‘utility’: Processing complexity and communicative efficiency. *Wiley Interdisciplinary Reviews: Cognitive Science*, *2*(3), 323-335.

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

Maurits, L. (2012). *Representation, information theory and basic word order* (Doctoral dissertation).

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

Regier, T., Kemp, C., & Kay, P. (2015). 11 Word Meanings across Languages Support Efficient Communication. *The handbook of language emergence*, *87*, 237.

Roberts, S. G., Torreira, F., & Levinson, S. C. (2015). The effects of processing and sequence organization on the timing of turn taking: a corpus study. *Frontiers in psychology*, *6*, 509.

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *language*, 696-735.

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

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

Shannon, C. (1948). A mathematical theory of communication. *Bell System Technical Journal*, *27*, 623–656.

Zipf, G. K. (1935). The psychology of language. *NY Houghton-Mifflin*.

Appendices

Appendix I – Demonstrating Informativeness

To clarify what 'informativeness' means, consider the following example:

Verb Informative: 'The man is painting a **______**.' Try to guess which word fits in the blank. The verb 'paint' is very restrictive, or predictive, of the upcoming Object Noun. In this case, the correct Object is 'picture'. Given the Verb 'paint', only a few notable guesses seem to be possible: picture, canvas, maybe house. Because the Verb conveys high information about what the upcoming Noun should be, the example statement is said to be Verb Informative.

Object Informative: 'The man eats **______**.' Try to guess which (specific) word fits in the blank. The verb 'eat' fits to a wide range of possibilities: fruit, salad, sandwiches, pasta, fish, candy, popcorn, etc. The correct Object for this example is 'bacon'. In the reverse direction, given 'The man **______** bacon,' only a few possible Verbs apply such as 'eats', 'fries', 'cooks'. Because the Object better restricts the possible Verbs than the other way around, statements like the initial example are said to be Object Informative.

Appendix II – Quantifying informativeness – Corpus analysis

For all corpus analysis in our study, we used the Corpus of Contemporary American English [COCA] (Davies, 2008). Since SOV and SVO orders both place the Subject at the beginning of communication, our experiments assume that the Subject is known and only looks at the relationship between the Object and the Verb. Our analysis is centered around probabilities used in Bayesian inference – the probability that a language producer uses an Object depends on how likely that Object is to co-occur with the given Verb (and vice-versa).

To quantify the relative informativeness between Objects and Verbs, we used COCA's collocate search function to find the collocate frequency of Objects given a Verb and vice versa. As an example, consider the Object-Verb pair "paint picture". We input the Verb (and associated word forms) "paint" into the search bar and looked for Nouns within 2 tokens following the word "paint" to account for words such as determiners that were not of interest for this study. Upon searching, COCA provides a list with the 100 most frequently co-occurring Nouns given the Verb "paint". To find the conditional probability of P(picture | paint), we divided the collocate frequency of "picture" by the sum of the top 100 collocate frequencies. To find the corresponding P(paint | picture), we repeated this process except looking for Verbs preceding the target noun within 2 tokens.

Appendix III – Conditional probabilities for Goldin-Meadow et al. (2008) stimuli

Verb Noun	**P(Verb | Object Noun)**	**Prob(Object Noun | Verb)**
Boy Stirs Spoon	0.0186	0.0035
Boy Tilts Glass	0	0
Woman Twists Knob	0.0427	0.0086
Girl Covers Box	0	0
Captain Swings Pail	0.0144	0
Man Picks Up Baby[1]	0	0.0069
Washer Woman Pets Dog	0.0033	0.1235
Man Plays Guitar	0.1367	0.0128
Man ***Moves*** Garbage ***Can*** to motorcycle man	0.0125	0
Man ***Carries***[2] ***Chicken*** to scaffolding	0	0
Dog ***Carries***[2] ***Flower*** to doghouse	0.0040	0
Bike With Cart ***Carries***[2] ***Girl*** to Giraffe	0	0

[1] No possible way to look up "Picks up" in COCA. 'Pick' also not seen as a result
[2] searched [carry]

Man ***Gives*** tool ***Box*** to Captain	0.0034	0
Girl ***Gives Flower*** to Man	0.0062	0
Man ***Throws Ball*** to basket	0.0193	0.0795
Workman ***Pushes Wheelbarrow*** to train	0.1240	0

Appendix IV – Stimulus Object-Verb pairs for our study

(Target word pairs in bold)

Noun informative

VO	P(V|O)	P(O|V)
Mix Cement	0.01929625	0.00595584
Mix Berries	**0.0250501**	**0.004357931**
Mix Bacon	0	0
Mix Cheese	0.01161314	0.009006392
Shut Drawers	0.01115619	0.0027358
Shut Curtains	**0.005407654**	**0**
Shut Bottle	0	0
Shut Bridge	0	0
Fill Bottle	**0.01078765**	**0.004644259**
Fill Bowl	0.02002298	0.008426697
Fill Curtains	0	0
Fill House	0	0.01139519
Eat Bacon	**0.02217997**	**0**
Eat Cheese	0.01596806	0.007519633
Eat Cement	0	0
Eat Berries	0.0068953	0
Build Bridge	**0.006190308**	**0.01343244**
Build House	0.02079898	0.05419639
Build Drawers	0	0
Build Bowl	0	0

Below is a less strict search for “Build Bridge” involving multiple forms. This search maintains the desired $P(V|O) > P(O|V)$ for the Noun Informative condition

P({Build, Building, Built}|Bridge*) = 0.108802
P({Bridges, Bridge}|Buil*) = 0.02463675

Verb Informative

VO	P(V|O)	P(O|V)
Wash Dish	**0.06776109**	**0.1198125**
Scrub Dish	0	0.001921062
Paint Dish	0	0
Watch Dish	0	0
Lift Weight	**0.02026402**	**0.05592352**
Carry Weight	0.02177408	0.04764618
Scrub Weight	0	0
Hang Weight	0	0
Watch TV	**0.1080882**	**0.1019373**
Turn TV	0.02714932	0.02239528
Wash TV	0	0
Sign TV	0	0
Paint Picture	**0.01917978**	**0.1570505**
Hang Picture	0	0.003130817
Type Picture	0	0
Lift Picture	0	0
Sign Contract	**0.104055**	**0.08552367**
Type Contract	0	0
Turn Contract	0	0
Carry Contract	0	0

Below are less strict searches for “Watch TV” and “Sign Contract” for which the above probabilities are not in the expected direction. Even with less strict searches, these pairs do not meet the criteria $P(V|O) < P(O|V)$ which is a potential confound for our study.

P({Watching, Watch, Watched, Watches}|TV) = 0.2280722
P({Watching, Watch, Watched, Watches}|Television) = 0.1877779

P({TV, Television}|Watch*) = 0.1658482

P({Signed, Sign, Signing, Signs}|Contract*) = 0.1331539

P({Contract, Contracts}|Sign*) = 0.03177691