No directional preference for grammaticalization in a semantic extension game

Anna Kapron-King[1,2], Simon Kirby[2], Graeme Trousdale[3] and Kenny Smith[2]
a.kapron-king@ed.ac.uk
[1]School of Informatics, University of Edinburgh
[2]Centre for Language Evolution, School of Philosophy, Psychology and Language Sciences, University of Edinburgh
[3]School of Philosophy, Psychology and Language Sciences, University of Edinburgh

Abstract

Grammaticalization is the process by which a lexical item (e.g., noun) acquires a more functional role (e.g., preposition) over time. Grammaticalization is considered largely unidirectional, that is, change from functional to lexical is far less common (Hopper & Traugott, 2003). What is the cause of this unidirectionality? Our experiment tests whether individuals have a preference in the direction of grammaticalization when performing semantic extension in communication. We focus on the phenomenon of using body part nouns as a source of spatial prepositions. We predicted that participants extending body parts to use as prepositions would find the task easier and more intuitive than participants extending prepositions to use as body parts. However, our results show no directional bias, indicating that the historical unidirectional tendency for body parts to be used as spatial relations does not originate in a bias that individuals have for using one to refer to the other.

Keywords: Semantic extension; grammaticalization; communication; unidirectionality; sender–receiver

Introduction

Grammaticalization is when a lexical item (a contentful item like a noun or verb) acquires a more grammatical function over time (becoming, e.g., a preposition or auxiliary). A pattern of grammaticalization that is widespread across languages is for terms referring to spatial relations to arise from body part nouns, for example, the English spatial term "ahead" meaning "in front (of)" comes from the body part noun "head". Grammaticalization has been widely observed in the world's languages, but change from grammatical to lexical (*degrammaticalization*) is much rarer and less systematic (Hopper & Traugott, 2003; Norde, 2009). An example of degrammaticalization is the Welsh verb *nôl* meaning "to fetch", which originated as the Middle Welsh preposition *yn ol* meaning "after" (Willis, 2007, as cited in Norde, 2009, pp. 148–150). The relative rarity of degrammaticalization is referred to as the *unidirectionality* of grammaticalization. Our aim is to understand where this unidirectionality comes from by studying what happens when participants perform semantic extension in either the direction of grammaticalization or degrammaticalization.

While grammaticalization takes place on long timescales, changes to the meanings of words are made possible by individuals extending the way they use those words in communication. Many grammaticalization researchers assume that individuals have asymmetric associations which affect how they extend the meanings of words, and this is what is responsible for unidirectionality of semantic change (e.g., Heine & Kuteva, 2011; Jäger & Rosenbach, 2008; Smith & Höfler, 2015). One reason for this assumption is the fact that metaphor is a key mechanism underlying semantic change, and is often considered asymmetric. According to Conceptual Metaphor Theory (CMT), metaphors allow us to talk about abstract concepts by relying on their associations with more concrete concepts (Lakoff & Johnson, 1980). This concrete–abstract asymmetry is widely credited as the cause of unidirectionality, as described by Heine and Kuteva (2011, p. 513):

> Underlying [grammaticalization] is a cognitive mechanism whereby concrete and salient concepts serve as vehicles or structural templates to conceptualize less concrete and less readily accessible concepts ... Thus, visible and tangible objects such as body parts or physical landmarks serve to express non-physical relations, such as spatial relations, and concrete actions serve as conceptual vehicles to express more abstract concepts describing the aspectual, temporal, or modal contours of events.

Compatible with this explanation, the Invited Inferencing Theory of Semantic Change (IITSC; Traugott & Dasher, 2001) says that the speaker introduces "innovative metaphorical relations" in conversation (Traugott & Dasher, 2001, p. 279). Hence, words that express concrete, external situations change to express the speaker's internal state, becoming more abstract (Traugott, 1989, pp. 34–35). The new meanings are inferred by the listener, and gradually spread to other language users.

Computational studies using datasets of semantic shifts have been used to test on a larger scale whether word meanings do in fact change from more concrete to more abstract in English (Xu, Malt, & Srinivasan, 2017) and cross-linguistically (Fugikawa et al., 2023; Winter & Srinivasan, 2022). Xu et al. (2017) and Fugikawa et al. (2023) found that concreteness was the best predictor of which concepts were used as source concepts historically, while Winter and Srinivasan (2022) found that frequency was better. However, Winter and Srinivasan (2022) analysed a small set of semantic shifts which only included nouns of similarly high concreteness, so this result does not provide much insight into unidirectionality of grammaticalization.

In D. Barner, N.R. Bramley, A. Ruggeri and C.M. Walker (Eds.), *Proceedings of the 47th Annual Conference of the Cognitive Science Society*

Studies like these are of course important in understanding patterns of grammaticalization, but they do not directly speak to the *cause* of these patterns. These studies are also limited by their reliance on lists of attested semantic shifts, which depend on synchronic colexification data for languages where diachronic data is unavailable, and rely on inferences that there was a change in a particular direction (Heine & Kuteva, 2002; Norde, 2009). This practice may augment the apparent prevalence of unidirectionality if counter-examples are overlooked because they would violate unidirectionality.

Experimental methods can be used to test proposed causes of unidirectionality, and do not suffer from the same concern regarding interpretational biases in natural language datasets of grammaticalization. Experiments have been used to show that the observed metaphoric asymmetry between space and time in language (i.e., spatial terms are used to refer to temporal concepts) is also present in people's non-linguistic mental representations of space and time (Boroditsky, 2000; Bottini & Casasanto, 2013; Casasanto & Boroditsky, 2008). However, there has yet to be work showing that this asymmetry affects how speakers engage in semantic extension (though Verhoef, Walker, & Marghetis, 2016 have studied extension in one direction, from space to time). Using a self-paced reading maze task, Hilpert and Correia Saavedra (2018) tested the "asymmetric priming hypothesis" (Jäger & Rosenbach, 2008), that unidirectional semantic change is caused by lexical items priming grammatical items, but not vice versa. They did not find the predicted effect, and in fact they found the opposite – reaction times were slower for the grammatical item when it was preceded by the related lexical item. Hilpert and Correia Saavedra (2020) tested the same hypothesis using a computational case study of the word "use", looking at whether corpus data can provide evidence of asymmetric priming, but again they found no support for it.

In this paper, we extend the work of Kapron-King, Kirby, Trousdale, and Smith (2024), who ran two artificial language experiments to test whether, when performing semantic extension, interlocutors are biased in favour of extending lexical items to refer to grammatical concepts and against extending the use of grammatical items to refer to lexical concepts. In their first experiment, participants judged the likelihood of artificial words being extended from body part to spatial preposition and from spatial preposition to body part. Participants' ratings were unaffected by the direction of extension. Their next experiment was a communication task, where participants learned artificial words for body parts or for spatial prepositions and then communicated with a partner about the opposite category, using only the words they had learned. Their results showed no difference between the two directions of extension in terms of communicative success, use of predicted extensions, or inter-participant agreement. However, their participants reported finding the task confusing and difficult, which may weaken these results.

Here we test the same hypothesis using a simpler design. One reason their communication task may have been difficult is that the participants first spent several minutes learning the meanings of new words and then they were unexpectedly required to extend them. In our experiment, participants do not learn a new vocabulary. Their experiment may also have been confusing because words were shown on their own, without a sentence for context. Textual context is considered to be key to the grammaticalization process (Hopper & Traugott, 2003). In our experiment, participants communicate using full sentences, which we expect will help them understand the nature of the task better. Our goal is to determine whether the results of Kapron-King et al. (2024), which suggest no asymmetry in extension consistent with a bias in favour of grammaticalization, continue to hold when the experimental design is simpler, more intuitive, and therefore more conducive to exemplifying asymmetries.

Methods

Our experiment is an imaginary communication task, where participants are told to pretend they are playing a communication game with a partner. Since we are studying initial extension biases, rather than the more long-term process of conventionalization, we would not anticipate a difference between how individuals behave with an imaginary partner and a real (online) partner.

Participants

We recruited 100 participants on Prolific, 25 in each condition. Participants were self-reported monolingual English speakers who had performed at least 50 prior experiments on Prolific with no less than a 99% approval rate. The experiment took 6 minutes to complete and participants were paid £1.14 (£11.44/hr). Participants provided informed consent before beginning.

Materials

The body parts and spatial preposition used as stimuli and responses in this experiment were derived from the *World Lexicon of Grammaticalization* (Kuteva et al., 2019), which contains a list of source-concept pairs of attested grammaticalization processes, based on data from over 1,000 languages. We found every instance of a body part concept grammaticalizing to a spatial concept, and converted these to mappings of English body part nouns and spatial prepositions[1]. This resulted in 23 body parts and 23 spatial prepositions, shown in Table 1. To ensure participants were aware they were performing extension between two different parts of speech, from noun to preposition or preposition to noun, we used full sentences as stimuli in this experiment. Senders were tasked with replacing the target word in a sentence, and receivers had to select a meaning (in the form of a sentence) given a sentence where the preposition or body part has been replaced with a word from the other category, see below.

[1] We include multi-word prepositional phrases, e.g., "on top of".

Table 1: Body parts and spatial prepositions used as stimuli.

body parts	back, belly, bowels, breast, buttocks, chest, ear, eye, face, flank, foot, forehead, guts, hand, head, heart, intestines, lip, liver, mouth, neck, shoulder, stomach
spatial prepositions	above, along, at, at the bottom of, at the edge of, behind, below, beneath, beside, in, in front of, in the middle of, inside, near, next to, on, on top of, out of, over top of, toward, under, upon, within

Procedure

Participants were told to pretend they were playing a communication game with a partner. They were assigned to one of two conditions: GRAM, meaning that they would be making extensions in the direction of grammaticalization, or DEGRAM, making extensions in the direction of degrammaticalization. They were also assigned to one of two roles: sender or receiver.

Each participant saw one trial for each of 23 stimuli. The order of trials was randomized. Examples of the sender and receiver trials are shown in Figure 1. Senders were given a sentence with one of the words bolded, then shown the same sentence with this word replaced by a blank space, and asked to fill in the blank by clicking one of 23 buttons[2].

Receivers were shown a sentence like the ones generated by senders, i.e., receivers in the GRAM condition saw "the fly is BODY-PART the box" and receivers in the DEGRAM condition saw "my PREP is sore". They were asked what the intended meaning of this sentence was, and given 23 buttons of sentences where the target item was replaced with one of the items from the other category.

Results

The results indicate no preference for extension from body part terms to prepositions compared to vice versa. We analysed the ease and intuitiveness of extension with three preregistered analyses: reaction times, consistency of responses, and use of predicted extension. Assuming that semantic extension is easier and more intuitive in the direction of grammaticalization than in the opposite direction, we expected faster reaction times, higher consistency of responses, and greater use of predicted extension in the GRAM condition than in the DEGRAM condition. All analyses were performed using the `brms` package (Bürkner, 2021) for Bayesian statistical analysis in R (R Core Team, 2024).

Condition was sum coded (GRAM $= -0.5$, DEGRAM $= 0.5$) and role was sum coded (sender $= -0.5$, receiver $= 0.5$).

[2]While it might be a more naturalistic task for participants to have to retrieve the body part or spatial preposition themselves by, e.g., typing it instead of clicking a button, we would have no way then to restrict their responses to the correct category.

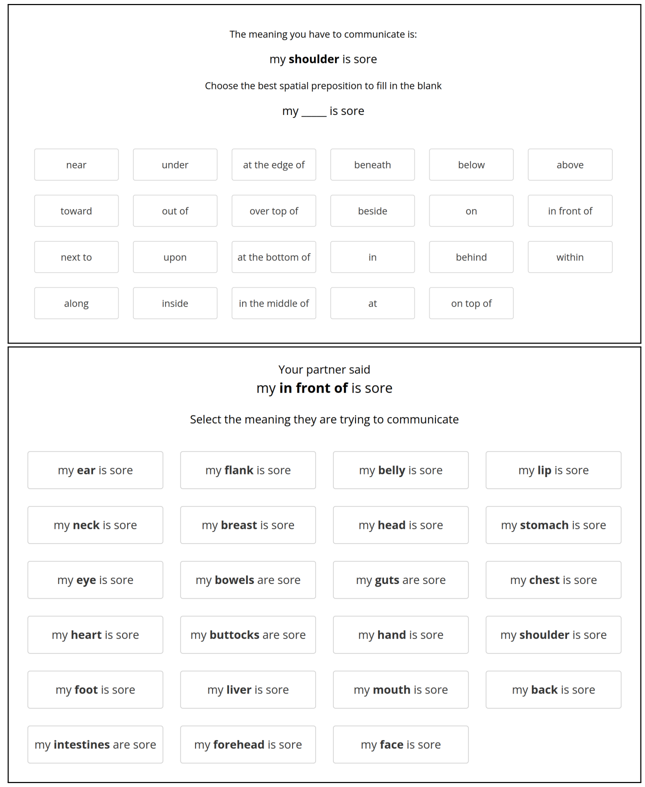


Figure 1: Screenshots of trials for a sender and receiver in the DEGRAM condition. Upper panel: Sender's view in the DEGRAM condition — the stimulus is "shoulder". Lower panel: Receiver's view in the DEGRAM condition — the stimulus is "in front of".

Responses that were reaction time outliers were excluded from all analyses. These were responses shorter than 400 ms and longer than 25000 ms. 196 responses out of 2300 were excluded, which included responses from 68 out of 100 participants.

Participant agreement

We measured participant agreement for a given stimulus using normalized entropy. The prediction was that participants in the GRAM condition would have stronger agreement (lower entropy) about how to extend body parts to spatial prepositions than those in the DEGRAM condition had about extending spatial prepositions to body parts. Normalized entropy h was computed as follows for each stimulus, where n was the total number of responses for that stimulus:

$$h = -\frac{\sum_{i=1} p_i \log(p_i)}{\log(n)}$$

This gives a value for h in the range (0,1), where 0 corresponds to complete agreement between participants (i.e., all participants select the same response), and 1 is maximal disagreement (participant responses are uniformly distributed over all possible responses). We fit a Bayesian linear model with a Beta likelihood and default priors. Because the data are aggregated over participants, we no longer have by-participant data. The formula for the linear model is:

$$h \sim \text{condition} * \text{role} + (1|\text{stimulus})$$

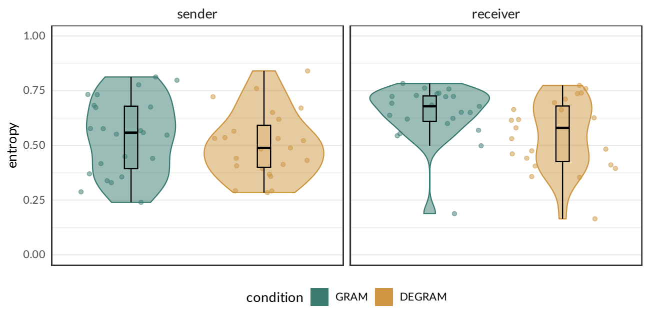
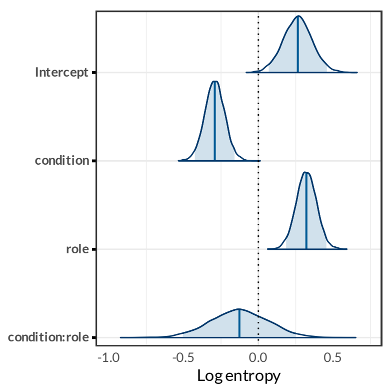
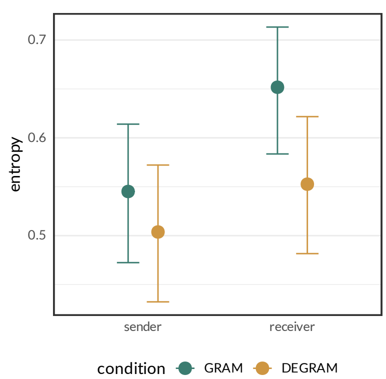


Figure 2: Entropy of responses. Each point is one stimulus. Higher entropy indicates less agreement between participants about how to refer to a given stimulus. Contrary to our prediction, entropy is lower in the DEGRAM condition.



(a) Posterior distributions for entropy model parameters. The dark vertical line under each curve indicates the mean β of the posterior and the shaded area is the 95% CrI.



(b) Conditional effects for entropy model.

Figure 3: Model results for entropy. Entropy is lower in the DEGRAM condition and higher for receivers.

The posterior distributions over parameters for the entropy model are shown in Figure 3a and conditional effect are shown in Figure 3b. The model estimates that there is a negative effect of condition ($\beta = -0.29$ log entropy, 95% CrI $[-0.43, -0.16]$), meaning that entropy is lower in the DEGRAM condition. In other words, contrary to our prediction, the model is confident that participants in the GRAM condition have less agreement in their responses than those in the DEGRAM condition. We also fit a null model which did not include condition as an effect. We compared it to the full model by computing the Bayes factor (BF), which is a ratio of the marginal likelihoods of two models. By making the null model's marginal likelihood the denominator and the full model the numerator, BF > 1 provides evidence for the full model and BF < 1 provides evidence for the null, with a greater magnitude indicating greater strength of the evidence (we use evidence categories from Lee & Wagenmakers, 2014). In this case the BF provides extreme evidence that the full model including `condition` and `condition:role` is better than the null model (mean $= 284.53$, sd $= 1.73$)[3].

The model also confidently estimates a positive effect of condition ($\beta = 0.32$ log entropy, 95% CrI $[0.19, 0.45]$), meaning that receivers have less agreement in their responses than senders.

There is a very uncertain negative estimate for the interaction ($\beta = -0.13$ log entropy, 95% CrI $[-0.51, 0.25]$), indicating a small possibility that the difference in entropy between senders and receivers is larger in the GRAM condition than in the DEGRAM condition.

Use of predicted extension

Responses were coded as predicted (1) or not predicted (0) according to whether the mapping between the body part and spatial preposition are present in the WLG (Kuteva et al., 2019). The proportions of predicted responses given by each participant are shown in Figure 4. We fit a Bayesian linear model with a Bernoulli likelihood distribution. The formula for the linear model is:

$$\text{pred} \sim \text{condition} * \text{role} + (1|\text{stimulus}) + (1|\text{participant\_id})$$

The posterior distributions over parameters for the predicted model are shown in Figure 5. The model estimates a slight positive effect of condition ($\beta = 0.12$ log odds, 95% CrI $[-0.21, 0.46]$), which would mean slightly more predicted responses in the GRAM condition than the DEGRAM condition, but the 95% CrI is close to being centred on zero, so there is not much certainty about this effect. The BF when comparing this model to a null model without condition indicates anecdotal evidence for the null hypothesis (mean $= 0.89$, sd $= 0.035$).

The model estimates a negative effect of role ($\beta = -0.75$ log odds, 95% CrI $[-1.09, -0.41]$), which means that receivers are less likely to use predicted responses than senders. This corresponds with the entropy results, indicating that the

[3] We used `brms::bayes_factor` which can have stability issues. All of ours were relatively stable. We report the mean and standard deviation of the BF after running ten times.

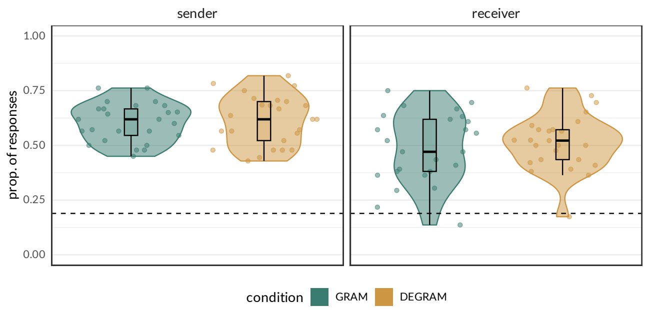


Figure 4: Proportion of predicted responses per participant. The dashed line shows the expected proportion of predicted responses for each participant. This is the mean number of responses that are predicted for each stimulus in a given condition (which happens to be the same in both conditions), divided by the total number of stimuli: 4.35/23. There is no difference between the GRAM and DEGRAM conditions.

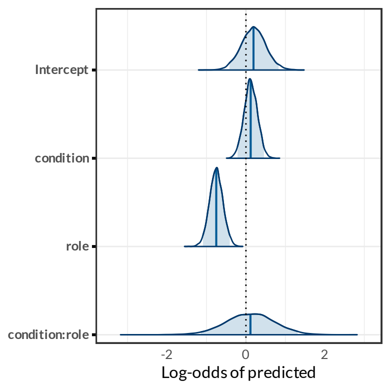


Figure 5: Posterior distributions for predicted responses model parameters.

task may be harder for receivers than senders. The model is not certain about the interaction effect ($\beta = 0.11$ log odds, 95% CrI $[-1.14, 1.31]$).

Reaction times

The reaction times are shown in Figure 7. In general, since participants have to choose between 23 possible responses on the screen for each trial, we did not expect reaction time to be very reliable for this task. We fit a Bayesian linear model with a log-normal likelihood and default priors. The formula for the linear model is:

$$\text{RT} \sim \text{condition} * \text{role} + (1|\text{stimulus}) + (1|\text{participant\_id})$$

The posterior distributions over parameters for the reaction time model are shown in Figure8. Rather than finding faster reaction times in the GRAM condition as we predicted, our results show the reverse effect. The model estimates a negative effect of condition ($\beta = -0.14$ log RT, 95% CrI $[-0.27, -0.02]$), indicating faster reaction times in the DEGRAM condition. However, the effect is quite small. The BF provided anecdotal evidence for the null hypothesis (mean $= 0.57$, sd $= 0.014$).

There is a slightly positive effect of role ($\beta = 0.04$ log RT, 95% CrI $[-0.08, 0.16]$), meaning that receivers may be a bit slower than senders, but the model is very uncertain about this. There is a slightly negative interaction effect ($\beta = -0.07$ log RT, 95% CrI $[-0.24, -0.10]$), which could mean a greater difference in reaction times between sender and receiver in the GRAM condition than the DEGRAM condition, but the model is very uncertain about this as well.

Exploratory

We used the `ggalluvial` package (Brunson, 2020) to visualize how participants responded for each stimulus, shown in Figure 6. This visualization made it clear that some prepositions (e.g., "out of", "near", "next to") were more challenging to choose a body part to express than others (in the GRAM condition), and some of the same prepositions were unlikely to be chosen as an extension for any of the body parts (in the DEGRAM condition). "Out of" was apparently the most difficult stimulus, because 5 data points were excluded due to being reaction time outliers. "Behind" and "buttocks" were the stimuli with the fastest mean response times in their respective categories. This is likely due to the fact that "behind" is a euphemistic synonym for "buttocks" in English.

In general, these results indicate that some mappings are more intuitive for participants, in both directions, which is likely affected by what associations are already present in their language and cultural context.

Discussion

Our results indicate that participants do not have a preference for performing semantic extension in the direction of grammaticalization compared to the direction of degrammaticalization. Participants in the GRAM condition did not show a greater use of predicted extensions, greater inter-participant agreement, or faster reaction times than those in the DEGRAM condition. If anything, we found slight evidence of an effect in the opposite direction, possibly because nominalization (using a word that is not a noun as a noun) is a familiar process for English speakers, while there is no equivalent process of "prepositionalization".

Unlike the results of Kapron-King et al. (2024), most of our participants did not seem to have very much trouble understanding the task. Comments we collected from participants tended to say they relied on their knowledge of where body parts are located in space, as we would expect, and this is borne out in the data. Senders in both conditions used the predicted extensions much more often than chance. Receivers found the task more difficult, as indicated by high entropy and low use of predicted extensions, but we had no predictions about how receivers differ from senders, given that the unidirectionality literature tends only to discuss the role of speakers. The absence of any certain interaction effects indicates that speakers and listeners engage with semantic extension between body parts and spatial prepositions in a similar

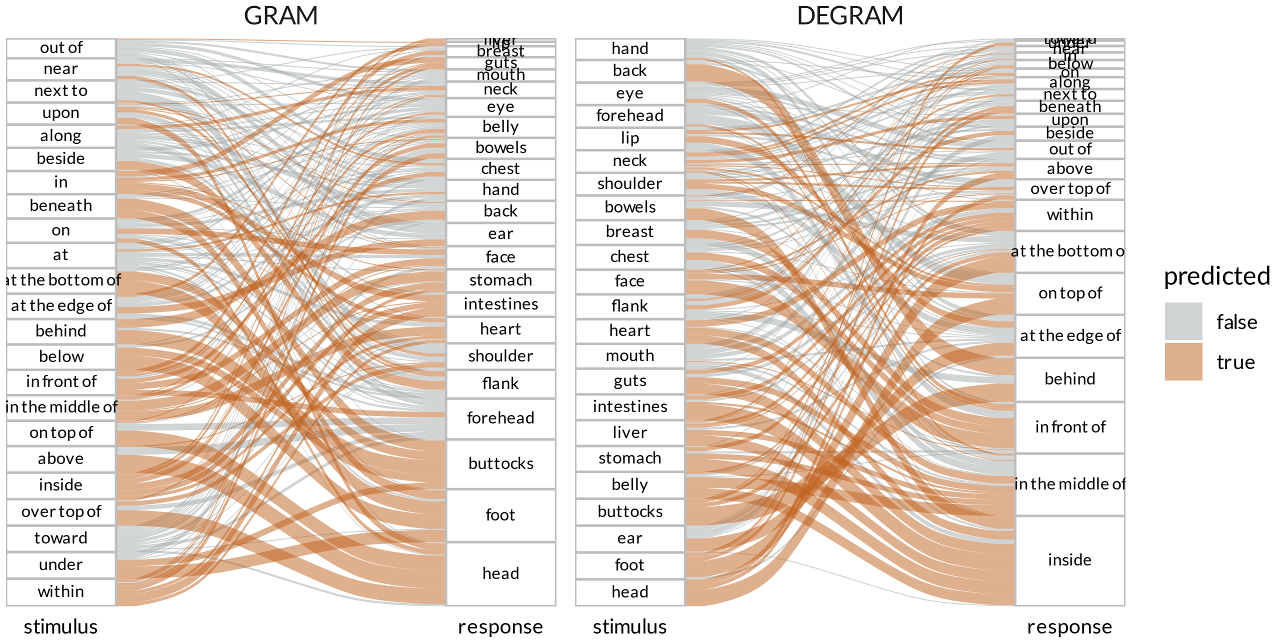


Figure 6: Graph showing, for senders in each condition, how often a certain response was produced for each of the 23 stimuli. The width of the line reflects the number of responses. Responses are in order from least used (top) to most used (bottom).

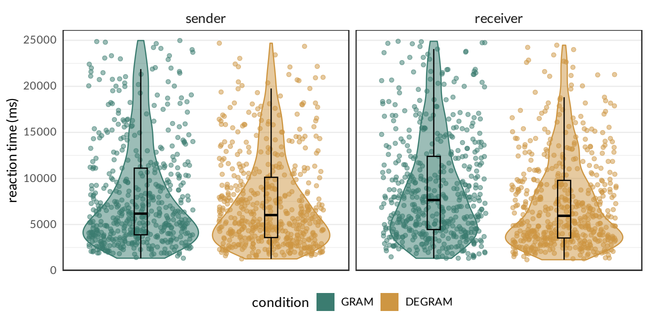


Figure 7: Reaction times for senders and receivers. Each point is one trial. Contrary to our prediction, participants in the DEGRAM condition responded faster than those in the GRAM condition.

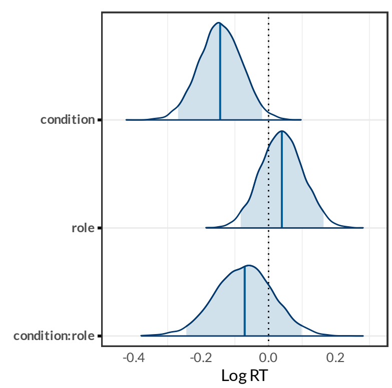


Figure 8: Posterior distributions of reaction time model parameters. In this figure, the intercept has been omitted for scaling purposes, because it is much larger than the other estimates ($\beta = 8.78$ log RT, 95% CrI $[8.70, 8.87]$).

way, with neither having any strong evidence of asymmetry.

Although our results are consistent with those of Kapron-King et al. (2024), we find it very surprising that we do not find evidence for such a widely held asymmetry assumption in the grammaticalization literature. This result provides further evidence that unidirectionality is not caused by a preference individuals have for extending lexical concepts to describe grammatical concepts. If, as so much of the research indicates, semantic change *does* exhibit unidirectionality, then the mechanism underlying this must be found at another stage of communication. It is possible that semantic extension in both directions is equally intuitive, but that lexical-to-grammatical extensions are more likely become conventionalized. Future research should investigate potential causes of unidirectionality in conventionalization.

Our work looks at body parts and spatial prepositions, but it may be that the metaphoric asymmetry between these domains has been overestimated. Future work should look at whether individuals behave asymmetrically when performing semantic extension between space and time, since, as we mentioned in the introduction, there is already a large body of evidence that people relate the concepts asymmetrically. Also, since the explanations of unidirectional grammatical change often rely on a more general assumption of asymmetric extension from *concrete* concepts to *abstract* concepts, it would also be worthwhile to test for this asymmetry when individuals perform semantic extension.

Acknowledgements

This project was supported by funding from UKRI (Grant Ref. EP/S022481/1).

References

Boroditsky, L. (2000, April). Metaphoric structuring: understanding time through spatial metaphors. *Cognition*, *75*(1), 1–28. doi: 10.1016/S0010-0277(99)00073-6

Bottini, R., & Casasanto, D. (2013). Space and time in the child's mind: metaphoric or atomic? *Frontiers in Psychology*, *4*.

Brunson, J. C. (2020). ggalluvial: Layered grammar for alluvial plots. *Journal of Open Source Software*, *5*(49), 2017. doi: 10.21105/joss.02017

Bürkner, P.-C. (2021). Bayesian item response modeling in R with brms and Stan. *Journal of Statistical Software*, *100*(5), 1–54. doi: 10.18637/jss.v100.i05

Casasanto, D., & Boroditsky, L. (2008, February). Time in the mind: Using space to think about time. *Cognition*, *106*(2), 579–593. doi: 10.1016/j.cognition.2007.03.004

Fugikawa, O., Hayman, O., Liu, R., Yu, L., Brochhagen, T., & Xu, Y. (2023). A computational analysis of crosslinguistic regularity in semantic change. *Frontiers in Communication*, *8*. doi: 10.3389/fcomm.2023.1136338

Heine, B., & Kuteva, T. (2002). *World lexicon of grammaticalization*. Cambridge, United Kingdom: Cambridge University Press.

Heine, B., & Kuteva, T. (2011, November). *Grammaticalization theory as a tool for reconstructing language evolution.* doi: 10.1093/oxfordhb/9780199541119.013.0054

Hilpert, M., & Correia Saavedra, D. (2018, November). The unidirectionality of semantic changes in grammaticalization: an experimental approach to the asymmetric priming hypothesis. *English Language & Linguistics*, *22*(3), 357–380. doi: 10.1017/S1360674316000496

Hilpert, M., & Correia Saavedra, D. (2020, October). Using token-based semantic vector spaces for corpus-linguistic analyses: From practical applications to tests of theoretical claims. *Corpus Linguistics and Linguistic Theory*, *16*(2), 393–424. doi: 10.1515/cllt-2017-0009

Hopper, P. J., & Traugott, E. C. (2003). *Grammaticalization* (2nd ed.). Cambridge University Press. doi: 10.1017/CBO9781139165525

Jäger, G., & Rosenbach, A. (2008). Priming and unidirectional language change. *Theoretical Linguistics*, *34*(2), 85–113.

Kapron-King, A., Kirby, S., Trousdale, G., & Smith, K. (2024). *Grammatical unidirectionality is not reflected in individual preferences when performing artificial semantic extension.* OSF Preprints. Retrieved from `osf.io/c5zks`

Kuteva, T., Heine, B., Hong, B., Long, H., Narrog, H., & Rhee, S. (2019). *World lexicon of grammaticalization* (2nd ed.). Cambridge, United Kingdom: Cambridge University Press.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago: University of Chicago Press.

Lee, M. D., & Wagenmakers, E.-J. (2014). *Bayesian Cognitive Modeling: A Practical Course* (1st ed.). Cambridge University Press. doi: 10.1017/CBO9781139087759

Norde, M. (2009). *Degrammaticalization*. Oxford ; New York: Oxford University Press.

R Core Team. (2024). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from `https://www.R-project.org/`

Smith, A. D. M., & Höfler, S. H. (2015, December). The pivotal role of metaphor in the evolution of human language. In J. E. Díaz-Vera (Ed.), *Metaphor and Metonymy across Time and Cultures* (pp. 123–140). doi: 10.1515/9783110335453.123

Traugott, E. C. (1989). On the rise of epistemic meanings in english: An example of subjectification in semantic change. *Language*, *65*(1), 31–55. doi: 10.2307/414841

Traugott, E. C., & Dasher, R. B. (2001). *Regularity in semantic change*. Cambridge University Press.

Verhoef, T., Walker, E., & Marghetis, T. (2016). Cognitive biases and social coordination in the emergence of temporal language. In *Proceedings of the 38th annual conference of the cognitive science society* (pp. 2615–2620). Austin, TX: Cognitive Science Society.

Willis, D. (2007, January). Syntactic lexicalization as a new type of degrammaticalization. *Linguistics*, *45*(2). doi: 10.1515/LING.2007.009

Winter, B., & Srinivasan, M. (2022, January). Why is semantic change asymmetric? the role of concreteness and word frequency and metaphor and metonymy. *Metaphor and Symbol*, *37*(1), 39–54. doi: 10.1080/10926488.2021.1945419

Xu, Y., Malt, B. C., & Srinivasan, M. (2017, August). Evolution of word meanings through metaphorical mapping: Systematicity over the past millennium. *Cognitive Psychology*, *96*, 41–53. doi: 10.1016/j.cogpsych.2017.05.005