Grammatical unidirectionality is not reflected in individual preferences when performing artificial semantic extension

Anna Kapron-King[1,2] (ID), Simon Kirby[2], Graeme Trousdale[3] and Kenny Smith[2]

[1]School of Informatics, University of Edinburgh, Edinburgh, UK; [2]Centre for Language Evolution, School of Philosophy, Psychology and Language Sciences, University of Edinburgh, Edinburgh, UK and [3]School of Philosophy, Psychology and Language Sciences, University of Edinburgh, Edinburgh, UK
Corresponding author: Anna Kapron-King; Email: a.kapron-king@ed.ac.uk

Abstract

Grammaticalization is the process whereby lexical items change into grammatical items. This phenomenon is widely attested, while the change from grammatical to lexical is far less common. We ran two experiments to test whether this unidirectional tendency originates with a preference for extending lexical meanings to grammatical ones rather than vice versa. We focus on body parts and spatial relations. In Experiment 1, participants were told the meaning of an artificial word then rated how likely it is that that word can also be used to refer to a second meaning – one meaning was a body part and one a preposition. We expected higher ratings when extending from body parts to prepositions than vice versa but found no difference. In Experiment 2, participants performed semantic extension in communication. We varied whether they extended words for body parts to prepositions or vice versa. Again, we found no asymmetry. Finally, we used a model of Experiment 2 to show that asymmetrical extension follows straightforwardly if there is an asymmetry in the number of words available relative to the number of meanings to express, indicating that having a larger number of lexical items than grammatical concepts could be an alternative source of unidirectionality.

Keywords: communication task; experimental; grammaticalization; semantic extension; unidirectionality

1. Introduction

Lexical items (e.g., nouns and verbs) provide meaningful content in an utterance, while grammatical items (e.g., pronouns and adpositions) provide grammatical structure to relate lexical items to each other. Often, these categories are considered to exist on a continuum, such that all linguistic units can be placed somewhere between fully lexical and fully grammatical (Haspelmath, 1999; Traugott & Trousdale, 2010).

When a lexical item acquires a more grammatical function, this is known as *grammaticalization* (coined by Meillet, 1912). A classic example is the grammaticalization of 'be going to' in English from a verb of motion in Old English, to a marker of intention around the 15th century, and finally to a future tense marker in the 18th century.

A widespread pattern of grammaticalization is the use of body part nouns as terms for spatial relations (e.g., spatial adpositions, locative adverbs) (Svorou, 1994). This phenomenon has been observed in many languages. For example in English, the spatial term 'ahead' (meaning 'in front (of)') comes from the body part noun 'head'; see examples (1) and (2) for further examples, and Table 1 for mappings of body part concepts to spatial concepts from the *World Lexicon of Grammaticalization* (2nd ed., Kuteva et al., 2019).

(1) 'stomach' → 'inside', Dagbani (Issah, 2020, p. 48)
    a. *M* **pùùní** *ká* *álááféé*
       my **stomach** has.NEG health
       'My stomach is paining (hurting) me'

    b. *Kúrígá* *máá* *tám* *dúú* *máá* **pùùní**
       bowl DEF stand DEF room **inside**
       'The bowl is inside the room'

(2) 'face' → 'on top of', Copala Trique (Hollenbach, 1995, pp. 174, 179, as cited in Kuteva et al., 2019, p. 169)
    a. **rian**^{32} *ne?e*^{3h} *a*^{32}
       face baby DEC
       'the baby's face'

Table 1. Observed instances of grammaticalization from body part to spatial concept (Kuteva et al., 2019)

Body part	Spatial concept
BACK	BEHIND
BACK	UP
BELLY	IN
BOWELS	IN
BREAST	FRONT
BUTTOCKS	BEHIND
BUTTOCKS	DOWN
EAR	LOCATIVE (e.g., 'region around the corner', 'at (the edge of)', 'towards')
EYE	FRONT
FACE	FRONT
FACE	UP
FLANK	SIDE
FOOT	DOWN
FOREHEAD	FRONT
HAND	LOCATIVE (e.g., 'in', 'at', 'from', 'into', 'through')
HEAD	FRONT
HEAD	UP
HEART	IN
LIP	LOCATIVE (e.g., 'edge', 'in', 'within', 'into', 'out of', 'at the edge of', 'along')
LIVER	LOCATIVE (e.g., 'in the middle', 'underneath', 'behind', 'inside')
MOUTH	FRONT
NECK	LOCATIVE (e.g., 'above', 'along', 'in')
SHOULDER	BEHIND
SHOULDER	UP

   b.   *oto*[32h]   *lu*[3]   **rian**[32]   *yana*[32!]   *a*[32]

      sleeps   cat   **face**   platform   DEC

      'The cat is sleeping on top of the platform'

While many regular patterns of grammaticalization have been observed across the world's languages, change in the other direction, from grammatical to lexical (*degrammaticalization*), is much rarer and less systematic (Hopper & Traugott, 2003; Norde, 2009). Thus, there is broad agreement that grammatical change has a strong unidirectional tendency.

Grammaticalization obviously involves syntactic and morphological changes (e.g., from content item to grammatical word to clitic to affix), but semantic change also plays an important role, especially in the early stages (Haspelmath, 1999; Heine, 2003; Hopper & Traugott, 2003; Sweetser, 1988; Traugott & Dasher, 2001). The semantic change component of grammaticalization is also thought to be largely unidirectional, and our aim is to understand where this unidirectionality comes from. We study this using behavioural experiments where participants perform semantic extension in either the direction of grammaticalization or degrammaticalization. In the following subsections, we present some proposed explanations of semantic unidirectionality (Section 1.1), describe previous empirical studies (Section 1.2), and provide background for our experiments (Section 1.3).

1.1. Explanations of semantic unidirectionality

A key assumption shared by explanations for unidirectionality in semantic change is that individuals have an asymmetric preference which affects how they extend the meanings of words. Metaphor is a major driver of semantic extension and change, and many agree with the conceptual metaphor theory (CMT; Lakoff & Johnson, 1980) view that metaphor is inherently asymmetric because it is grounded in experience. For example, the metaphors MORE IS UP and LESS IS DOWN, as in 'crime rates are falling', come from the experience of observing that piles with more items in them are higher. Hence, metaphors allow us to talk about abstract concepts by relying on their associations with concrete concepts (Lakoff & Johnson, 1980). This is commonly considered the cause of unidirectionality, as described by Heine and Kuteva (2012, 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.

On its own, CMT does not provide a specific explanation for how asymmetric metaphor leads to unidirectional change. A potential linking mechanism of asymmetric priming was proposed by Jäger and Rosenbach (2008). For example, if the motion verb *be going to* evokes (i.e., primes) the related concept of intention, the speaker is more likely to talk about intention in the following discourse and they are more likely to use *be going to*. Over time, this would lead to the intention meaning of

be going to becoming conventionalized. If priming is asymmetric, that is, the motion term primes the intention term more than the reverse. This is more likely to happen in the motion-to-intention direction than vice versa, leading to unidirectional change (Jäger & Rosenbach, 2008, pp. 104–106).

An alternative explanation, which does not depend on CMT (yet is not incompatible with it), comes from the Invited Inferencing Theory of Semantic Change (IITSC) (Traugott & Dasher, 2001). IITSC says that semantic change begins when a speaker uses a word in a new way, with the new meaning being conveyed through implicature. The listener may then infer the speaker's intended meaning. Through repeated inferencing, the new meaning becomes conventionalized, leading to semantic change. Because the speaker is the primary agent of change, IITSC predicts that *subjectification* is the main type of semantic change – that meanings change to express the point of view of the speaker over time. Consequently, words that express '[m]eanings based in the external described situation' change to express 'meanings based in the internal (evaluative/perceptual/cognitive) described situation' (Traugott, 1989, pp. 34–35), resulting in unidirectionality.

Haspelmath (1999) claims that the invisible-hand theory of language change (Keller, 1989) can explain unidirectionality. This theory says that individuals take linguistic actions which follow certain maxims, naturally leading to change. According to Haspelmath (1999), the maxim of clarity prevents speakers from using grammatical items as lexical items because grammatical items are 'less salient and less explicit than lexical items' (p. 1059). While the cause of the asymmetry is quite different from the other theories, it still posits an asymmetric preference on the part of the speaker.

1.2. Previous empirical studies

The above explanations relied on case studies and small corpus studies as evidence. More recently, large-scale corpus studies and behavioural experiments have been applied to unidirectionality of semantic change.

Large-scale studies of semantic shifts have explored what linguistic factors best characterize historical patterns of semantic change in English (Xu et al., 2017) and cross-linguistically (Fugikawa et al., 2023; Winter & Srinivasan, 2022) but have produced conflicting results. These studies used existing lists of semantic shifts, represented as source–target concept pairs and applied computational methods to find what best predicts the directionality of the shifts. Xu et al. (2017) found that externality and embodiment (both facets of concreteness) were the best predictors for semantic shifts in English, consistent with the CMT-based explanation that semantic change goes from concrete to abstract. Fugikawa et al. (2023) provide a similar result using a cross-linguistic dataset. In a separate corpus, Winter and Srinivasan (2022) found that frequency was a better predictor of directionality than concreteness, though the concepts analyzed in this study were all fairly concrete nouns.

Large-scale corpus studies are of course essential to test whether patterns found in smaller datasets hold more generally, but they are necessarily correlational and cannot examine the *cause* of these patterns. These corpus studies are also limited by their reliance on lists of attested semantic shifts. These lists often rely on synchronic data for many of the languages that do not have diachronic sources, rely on inferred changes (Heine & Kuteva, 2002; Norde, 2009). This may augment the

apparent prevalence of unidirectionality, if counterexamples are overlooked because they would violate assumed 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. Hilpert and Correia Saavedra (2018) tested Jäger and Rosenbach's asymmetric priming hypothesis using a 'maze task', a variant of self-paced reading where participants construct a sentence one word at a time. The sentences contained a lexical and grammatical use of the same word (see example 3b).

(3) a. The student kept$_{\text{lexical}}$ the light on to keep$_{\text{grammatical}}$ reading.
     b. The student kept$_{\text{grammatical}}$ checking facebook to keep$_{\text{lexical}}$ up to date.

They predicted that if lexical items prime grammatical items but not vice versa, reaction times to the second use would be faster when a lexical use preceded a grammatical use (3a), but there would be no effect of priming for the opposite order (3b). Their results did not show the predicted effect, and in fact, they found that 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 the word 'use' as a case study, looking at whether corpus data can provide evidence of asymmetric priming, but again found no support for it.

Various experiments have shown that the widespread metaphorical asymmetry between space and time in language is based in a more fundamental asymmetry in people's mental representations of space and time (Boroditsky, 2000; Bottini & Casasanto, 2013; Casasanto & Boroditsky, 2008), consistent with the idea that such asymmetries may be at the root of asymmetries in semantic extension and therefore unidirectionality. However, while Verhoef et al. (2016) have studied extension from space to time, no work to date has tested extension in the reverse direction, a necessary step to establish that the asymmetry is present in semantic extension.

1.3. Our experiments

The common assumption in the literature reviewed above is that speakers prefer extension in the direction of grammaticalization rather than degrammaticalization. Our hypothesis is therefore that when performing semantic extension, interlocutors are biased in favour of extending the use of lexical items to refer to grammatical concepts and against extending the use of grammatical concepts to refer to lexical concepts. This bias then leads to unidirectionality in how the meanings of words change over time. We tested this hypothesis using two artificial language experiments in which participants engaged in semantic extension. Our experiments involve extension between body parts and spatial relations, as this is often cited as a straightforward example of unidirectionality in grammaticalization research.

In Experiment 1, participants rated the likeliness of a given semantic extension between a body part and spatial preposition. Likeliness ratings are a common way of measuring participants' instincts about language in psycholinguistics research, although participants are usually asked to judge the acceptability of entire sentences (e.g., Suttle & Goldberg, 2011).

In Experiment 2, we study semantic extension in communicative interaction between pairs of participants. Artificial language learning and communication

experiments have been widely used to study whether biases acting in learning and communication are responsible for universals of language (e.g., Bowerman & Smith, 2022; Culbertson et al., 2012; Kanwal et al., 2017; Karjus et al., 2021). Our Experiment 2 adapts methods from the sender–receiver task used by Karjus et al. (2021) and Bowerman and Smith (2022) to study semantic extension: participants learn artificial words for body parts or prepositions and are then forced to extend those terms to to communicate with their partner about concepts from the other domain (e.g., attempting to extend body part terms to convey prepositional concepts).

2. Experiment 1: Individual judgement task

In this experiment, we test the hypothesis that participants are more likely to accept a word introduced as a label for a body part being used as a spatial preposition than a word introduced as a spatial preposition being used as a body part. We ran an artificial language task with two conditions, in a within-subjects design, where participants rated the likeliness of a body part term being extended to a spatial preposition or vice versa.

2.1. Methods

2.1.1. Participants

Participants were 200 self-reported monolingual English speakers, recruited via Prolific. We chose monolingual English speakers to control for any effect that knowledge of another language might have on participants' judgements about the artificial language. The experiment took just over 2 minutes to complete on average, and participants were paid £0.48.

2.1.2. Materials

On each trial, participants were given the meaning of an artificial word and asked to judge how likely it is that that word could also be used to refer to a second meaning (see Figure 1). The stimuli for each trial consisted of an artificial word and two meanings: a body part and a spatial preposition. The pairs of body parts and spatial prepositions were derived from Heine and Kuteva's *World Lexicon of Grammaticalization* (1st ed., 2002). This book contains observed instances of grammaticalization from over 1,000 languages in the form of a 'source–target lexicon'. We found all entries where the source concept was a body part and the target concept was a spatial relation, for example, BOWELS ('bowels', 'guts', 'intestines') → IN (SPATIAL) (p. 82), then chose one English body part noun and one spatial preposition to represent this pair of concepts (in this case, 'guts' and 'in'). We did not use any spatial prepositions that already contained a body part, such as 'ahead' and 'in back of'. This meant that some concept pairs had to be excluded, since there was no suitable choice of spatial preposition. The resulting list of 20 pairs is shown in Table 2.

Each participant completed 10 trials, encountering 10 randomly selected pairs from this list once each. Because some body parts and spatial prepositions appear in our list more than once, pairs were selected such that participants never saw the same body part or the same preposition twice (e.g., no participant saw both buttocks–under and foot–under, or both face–on and face–in front of).

Figure 1. Example of a SPACE-TO-BODY trial from Experiment 1. The participant saw the first screen, then gave their rating on the second screen.

Table 2. The pairs of meanings used as stimuli in Experiment 1, derived from Heine and Kuteva (2002)

Body part	Spatial preposition
buttocks	under
breast	in front of
chest	near
ear	at the edge of
eye	in front of
face	on
face	in front of
flank	next to
foot	under
forehead	in front of
guts	in
head	above
head	in front of
heart	in
lip	along
liver	in the middle of
mouth	in front of
neck	above
shoulder	up
stomach	within

We generated 20 CVCV artificial words to use as labels: *reva, viku, havi, vipa, rapi, melu, vamu, pevo, kapi, neto, voki, tuta, pona, nehu, lavo, tiro, lapo, mero, meti, nulu*. Two syllables is a plausible length in English for both nouns and prepositions, with the intention being to avoid biasing participants towards particular responses based purely on word length. We selected 10 words at random for each participant.

2.1.3. Procedure

This experiment was made using jsPsych (de Leeuw et al., 2023) and ran in the participant's browser. Participants were told that they would be making judgements about words from a language called Aki, spoken by 5,000 people on an island in the Pacific. This deception was intended to encourage participants to respond as they would for a real language, following e.g., Saldana et al. (2021). Participants were debriefed on the deception and the purpose of the experiment upon completion.

In each trial, participants were told the meaning of an artificial word and then on the next page were asked to rate how likely it is that that same word can also be used to mean a different meaning. Participants responded using a continuous sliding scale from *very unlikely* to *very likely*. The responses were recorded as values in the range [0,1]. The use of two separate screens and the wording of the question ('can also be used to mean') were intended to imply a direction of extension.

There were two conditions in a within-subjects design. In five trials, the first meaning given for the artificial word was a body part, and the second was the corresponding spatial preposition (BODY-TO-SPACE condition). In the other 5, the first meaning was a preposition, and the second was the corresponding body part (SPACE-TO-BODY condition). Participants encountered the two trial types in random order.

2.1.4. Analysis

To determine whether there is an asymmetry in participants' willingness to extend a body part to a spatial preposition compared to vice versa, we compared the ratings produced in the two different conditions. Our prediction was that participants would produce a higher rating in the BODY-TO-SPACE condition than in the SPACE-TO-BODY condition. This result would provide evidence for the claim that unidirectionality on a historical scale originates from an asymmetry in how individuals engage in semantic extension between lexical and grammatical items. This analysis was pre-registered on OSF, and the data and analysis code are also available there.[1]

2.2. Results

The results are shown in Figure 2. There is no obvious difference in the ratings between the two experimental conditions, indicating that participants do not feel differently about a body part term being used as a spatial preposition than they feel about a spatial preposition being used as a body part. The lack of asymmetry also seems to hold across all stimuli pairs, as shown in Figure 3.

To verify this result, we fit a mixed-effects beta-regression in R version 4.1.1 (R Core Team, 2023) using the glmmTMB package (Brooks et al., 2017). The dependent variable was the rating, and the independent variable was the experimental condition. Beta-regression only allows for response variables within (0,1), i.e., excluding the extremes of 0 and 1, but our scale included 0s and 1s. We fit both a zero-inflated model (where 0 is modelled separately from the rest of the responses) with 1 responses capped to 0.99, and a regular beta-regression model with responses transformed to be within the boundaries. We performed the transformation suggested by Smithson and Verkuilen (2006), using the formula $y = (y * (N - 1) + 1/2)/N$, where N is the number of participants. This resulted in responses being in the range [0.0025,0.9975].

---

[1]Preregistration: https://osf.io/t7yr8. Data and analysis code: https://osf.io/kzs7u.
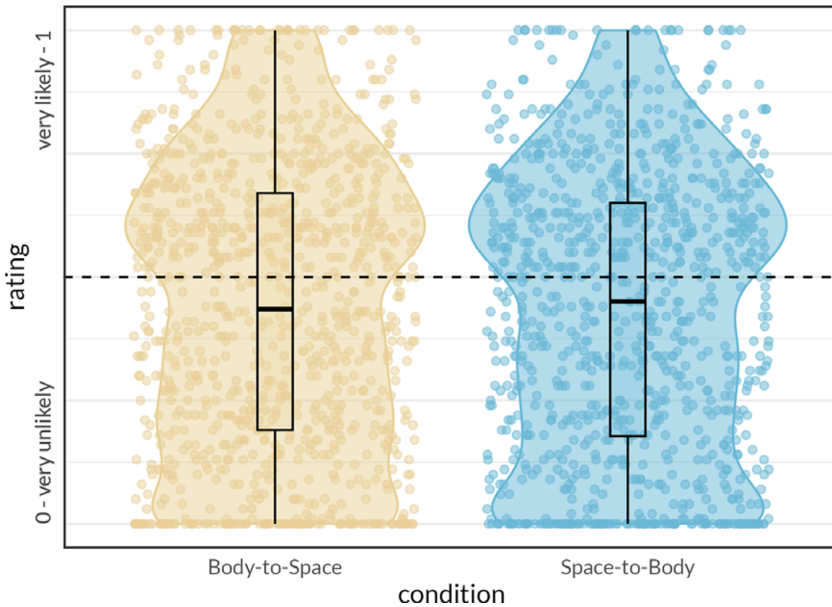
Figure 2. Experiment 1 ratings. Each dot represents a single response from a single participant. Response values are on a scale from *very unlikely* to *very likely*. The responses are split by condition: BODY-TO-SPACE, where the first meaning was a body part and the second was a preposition, and SPACE-TO-BODY, where the order was reversed. The violin plots show the density of responses along the y-axis. The box plots indicate the 25th percentile (lower hinge) and 75th percentile (upper hinge) with a dark line indicating the median. The whiskers extend $1.5 * IQR$ from the hinges. Contrary to our prediction, there was no difference in participants' responses between the two conditions.

Transforming the responses in this way may be reasonable because the ends of our scale (*very unlikely* and *very likely*) were not worded absolutely and thus do not necessarily need to be modelled separately. The model formulas, including random effects, are shown in Table 3.[2]

There was no effect of condition for either the zero-inflated beta-regression model ($\beta = -0.073$, SE $= 0.046$, $z = -1.59$, $p = 0.11$, ZI: $\beta = 0.26$, SE $= 0.28$, $z = 0.94$, $p = 0.35$) or the standard model with transformed data ($\beta = -0.080$, SE $= 0.052$, $z = -1.53$, $p = 0.13$).

2.3. Discussion

Unexpectedly, we found no preference for extending body parts to spatial prepositions over extending spatial prepositions to body parts, and therefore, no evidence supporting the widely assumed asymmetry in how speakers engage in semantic extension that is taken to account for unidirectionality in grammaticalization.

---

[2]For both the standard and zero-inflated models, we tried fitting the maximal model, `response ~ condition + (condition | participant_id) + (condition | pair)` but it did not converge for the zero-inflated model, so here we report the most complex models that did converge.
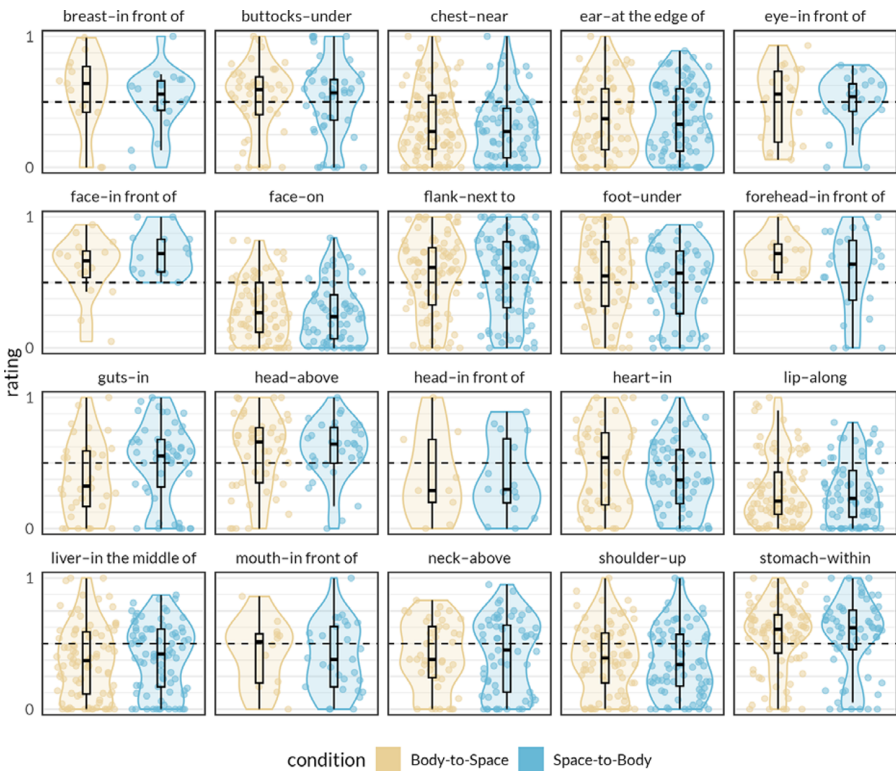
Figure 3. Experiment 1 ratings by body part–preposition pair, showing no clear preference for the predicted direction of semantic change. Plotting conventions as in Figure 2.

Table 3. Formulas for the mixed-effects beta-regression models fit to the Experiment 1 data

	Zero-inflated beta model
Formula ZI	response ~ condition + (condition | participant_id) + (1 | pair) response ~ condition + (1 | participant_id) + (condition | pair)
	beta model with transformed data
Formula	response ~ condition + (condition | participant_id) + (condition | pair)

It is of course possible that this null result is due to the design of our task. It may not have been clear to participants that the first meaning was the established meaning and the second meaning was a potential extension; participants may have been responding with how likely they thought it was that a word could have both meanings at the same time, rather than considering an extension in meaning. Alternatively, the lack of asymmetry could be because participants were not engaging in a communicative task but merely judging the acceptability of extensions. Individuals producing extensions in genuine interaction are required to consider how understandable a potential extension would be to an interlocutor, and this estimation of interpretability may be a crucial mechanism for the expected asymmetry.

3. Experiment 2: Communication task

Experiment 2 addresses some of the potential flaws in Experiment 1 by (1) making the semantic extension nature of the task more obvious and (2) having participants make semantic extensions in communication.

3.1. Methods

3.1.1. Participants

Participants were 243 monolingual English speakers recruited via Prolific. The data from 200 participants are included in our analysis, with the remaining participants lost due to failure to pair with a partner for the communication phase (see below). The experiment took approximately 30 minutes, and participants were paid £4.80 (£9.60/h).

3.1.2. Materials

Using the body part and preposition meanings from Experiment 1, we generated four lists of six pairs of body parts and prepositions, under two constraints: the same body part or preposition never appeared more than once in each list; lists did not contain multiple body parts that could map to near-synonymous prepositions (e.g., we would not accept a list that had 'heart' and 'stomach', because they both map to 'in' and 'within'). Our intention was to have as little confusion as possible about which body parts and prepositions could be associated with each other in the communicative task. The four lists are shown in Table 4.

We used the same two-syllable artificial words from Experiment 1, and each participant was given their own random selection of six words (see below). Participants were trained on one of the four lists, and the 200 participants were fairly evenly spread between the lists (list 1: 54 participants, list 2: 54, list 3: 46, list 4: 46).

3.1.3. Procedure

The experiment was implemented in jsPsych, run on top of a Python server to enable participants to interact using WebSockets.[3] In the training phase, participants were trained on an artificial vocabulary of six body parts (BODY-TO-SPACE condition) or six prepositions (SPACE-TO-BODY condition), with meanings coming from one of the four lists in Table 4. For example, a participant in the BODY-TO-SPACE condition with stimuli from list 1 would learn artificial words meaning 'heart', 'lip', 'face', 'foot', 'chest', and 'mouth'. The training phase included three kinds of trial: observation, comprehension, and production (see Figure 4). In the observation trials, participants were simply presented with the meaning of the artificial word for 3 seconds. On comprehension trials, participants selected the meaning of a word from an array of options. For production trials, participants were given a meaning and had to select the corresponding word from an array. Participants were always given the same array of six meanings or six words to choose from on comprehension and production trials (with the order held constant) and received feedback after each comprehension and production trial, telling them if they were correct, and what the correct response was

---

[3]The code is a modification of the materials written by Kenny Smith for the University of Edinburgh course Online Experiments for Language Scientists: https://kennysmithed.github.io/oels2021/oels_practical_wk10.html.

Table 4. Lists of body parts and spatial prepositions used as stimuli in Experiment 2

	Body parts (seen targets for BODY-TO-SPACE)	Spatial prepositions (seen targets for SPACE-TO-BODY)
1	heart, lip, face, foot, chest, mouth	in, along, on, under, near, in front of
2	lip, buttocks, mouth, neck, flank, liver	along, under, in front of, above, next to, in the middle of
3	guts, lip, face, neck, flank, shoulder	in, along, on, above, next to, up
4	heart, lip, chest, forehead, neck, ear	in, along, near, in front of, above, at the edge of
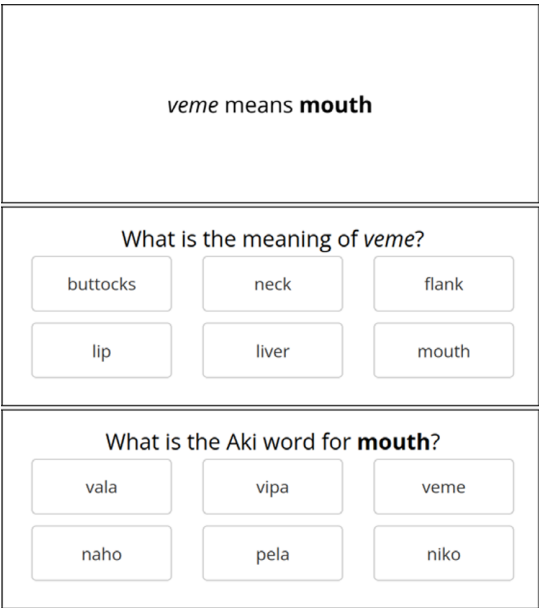
Figure 4. The three kinds of training used in the first phase of Experiment 2. Upper panel: an observation trial, the participant passively observes the word plus associated meaning for 3 seconds. Middle panel: comprehension trial, the participant selects the meaning of a word and receives feedback. Lower panel: production trial, the participant selects the word for a given meaning and receives feedback.

if they were wrong. Training consisted of four blocks, each block composed of six observation trials, six comprehension trials, and six production trials in that order, covering each of the six word-meaning pairings once in each trial type, for a total of 72 training trials.

After training, participants were sent to a virtual waiting room for a maximum of 7 minutes while they waited to be automatically assigned a partner. Participants were paired with another participant who was in the same condition and had been trained on the same six meanings (but not necessarily the same six word forms, see below). There were 22 participants who were unable to start the communication phase due to a lack of an available partner. A further 20 participants entered but did not complete the communication phase, due to one participant dropping out. Non-completion was considered as withdrawal of consent and all of the data from participants who did not make it to the end of the experiment was therefore excluded from our analysis.

Additionally, one participant was excluded because their partner participated in the experiment twice; for the repeat participant, only the data from their first participation was included in our analysis.

Once paired, participants began the communication phase, where they took turns playing the roles of sender and receiver (see Figure 5). The sender's task was to help the receiver identify the target meaning from an array of meanings by choosing a word to send from the six words they were trained on. The receiver was then given the word the sender had chosen and had to select the correct meaning. Both partners saw the same array of six meanings, which were independently randomised in order. After the receiver responded, both partners were given feedback indicating whether the receiver was correct or incorrect, as well as what the target meaning was and what meaning the receiver had selected. Because participants each learned a different random assignment of artificial words to meanings, the word the receiver saw was not really the word the sender selected, but instead the word from the receiver's lexicon that corresponded to the same meaning in training (e.g., if during training the sender had seen 'veme' paired with mouth, and the receiver had seen 'naho' paired with mouth, then whenever the sender selected 'veme' the receiver would see 'naho'). A similar remapping procedure is used in Silvey et al. (2015) and is intended to reduce

Figure 5. The two roles in the communication phase of Experiment 2. Upper panel: the sender's view, the participant chooses a word to refer to the highlighted meaning. Lower panel: the receiver's view, the participant sees a word and selects the intended meaning.

any preferences participants may have to associate certain meanings with certain forms, e.g., on the basis of iconic sound-meaning correspondences.

Since the highlighted target meaning for the sender is a body part ('chest'), this serves either as an example of the BODY-TO-SPACE condition for a seen target, or the SPACE-TO-BODY condition for an unseen target.

In total, there were 60 trials in the communication phase. In the first 12 trials, participants communicated about the meanings they learned in training (*seen targets*), both acting as sender once for each of the six meanings. This initial block of trials was meant to help them understand the communicative task. After that, trials with *unseen targets* were included. Unseen targets were meanings from the same list that the two partners had been trained on, but from the opposite semantic domain. For example, the unseen targets for partners in the BODY-TO-SPACE condition with stimuli from list 1 would be 'in', 'along', 'on', 'under', 'near', and 'in font of'. Both partners acted as the sender twice for each unseen meaning, and a further two times for seen meanings.

3.1.4. Analyses

We analysed the data from the communication phase, looking at whether there was an asymmetry in how participants in the BODY-TO-SPACE and SPACE-TO-BODY conditions behaved for unseen targets compared to seen targets. For seen targets, we expected no difference between the two conditions. For unseen targets, in the BODY-TO-SPACE condition we expected higher communicative success, greater use of the predicted response (i.e., extending according to the frequent grammaticalizations seen in Heine and Kuteva (2002), e.g., extending the word for 'mouth' to mean 'in front of', and doing so more reliably than the word for 'in front of' was extended to mean 'mouth'), and less variation in how senders responded for a given target meaning.

We capture these three predictions with three different analyses. Communicative success was measured by whether or not the receiver responded with the correct meaning based on the sender's chosen word, and was analysed using logistic regression, using the lmerTest package (Kuznetsova et al., 2017). The predicted response for a given target depended on whether the target meaning was seen or unseen. For seen targets, the predicted response was trivially the word that the participant had learned for the target meaning. For unseen targets, it was the word whose meaning was paired with the target meaning in Table 2, i.e., the frequent pairing from the World Lexicon of Grammaticalization; again, this is binary data which we analyse with logistic regression. Finally, to measure how varied the senders' responses were, we computed the conditional entropy of responses for each target meaning, with higher entropy corresponding to more varied label choices for a given meaning. We expected lower entropy when participants were extending body parts to express spatial terms in the BODY-TO-SPACE condition, indicating that there is more agreement between senders about which body part to use to refer to a given spatial preposition, relative to the agreement about how to extend spatial prepositions to refer to body parts. Entropy data was analysed using linear regression. The preregistered analyses (Sections 3.2.1, 3.2.2, and 3.2.3), data and analysis code are available on OSF.[4] Summaries of the fixed effects for all analyses are in the Supplementary Materials.

---

[4]Preregistration: https://osf.io/5pd8u. Data and analysis code: https://osf.io/zukya.
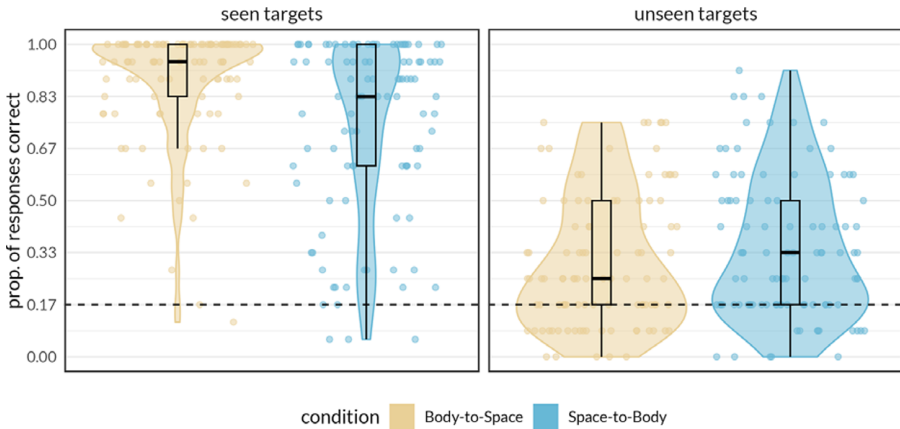
Figure 6. Communicative success in Experiment 2. Each dot represents a single participant and indicates the proportion of their responses as the receiver that were correct. The dotted line indicates chance performance, i.e., random selection from among 6 possible meanings. The 'seen targets' facet shows results for trials where the target was one of the six meanings the participant encountered during training: body parts for the BODY-TO-SPACE participants and prepositions for the SPACE-TO-BODY participants. The 'unseen targets' facet shows results for trials where the target was from the opposite category from what the participant saw in training. As expected, receivers are highly successful for seen targets. However, unexpectedly, success is lower when the seen target is a preposition, suggesting these may be harder to learn. We expected receivers to have higher correctness for unseen targets in the BODY-TO-SPACE than the SPACE-TO-BODY condition, but found no significant difference.

### 3.2. Results

3.2.1. Communicative success

Communicative success is shown in Figure 6. The logistic regression included fixed effects of condition (sum coded), seen (treatment coded), and their interaction, with a random intercept for target meaning, and a nested random intercept for pair and participant.[5] There was no effect of condition on communicative success ($\beta = -0.13$, SE $= 0.14$, $z = -0.89$, $p = 0.38$), though there was a significant interaction between condition and seen ($\beta = 2.90$, SE $= 0.08$, $z = 35.45$, $p < 0.001$). While there was no difference between the two conditions for unseen targets, SPACE-TO-BODY participants unexpectedly performed worse than BODY-TO-SPACE participants for seen targets. The difference in performance for seen targets could be because new prepositions were harder to learn than body parts.[6]

3.2.2. Predicted extensions

Use of predicted responses is shown in Figure 7. Again there was also no effect of condition ($\beta = -0.018$, SE $= 0.158$, $z = -0.11$, $p = 0.91$), and there was a marginally

---

[5]Including random slopes for seen led to failure to converge.

[6]Further evidence for this comes from analysing participants' performance on the training trials, which showed that BODY-TO-SPACE participants performed slightly but significantly better on both comprehension trials ($p < 0.001$) and production trials ($p < 0.01$) than SPACE-TO-BODY participants. Mean proportion correct on comprehension trials was 0.80 for BODY-TO-SPACE participants and 0.72 for SPACE-TO-BODY. Mean proportion correct on production trials was 0.84 for BODY-TO-SPACE participants and 0.76 for SPACE-TO-BODY.
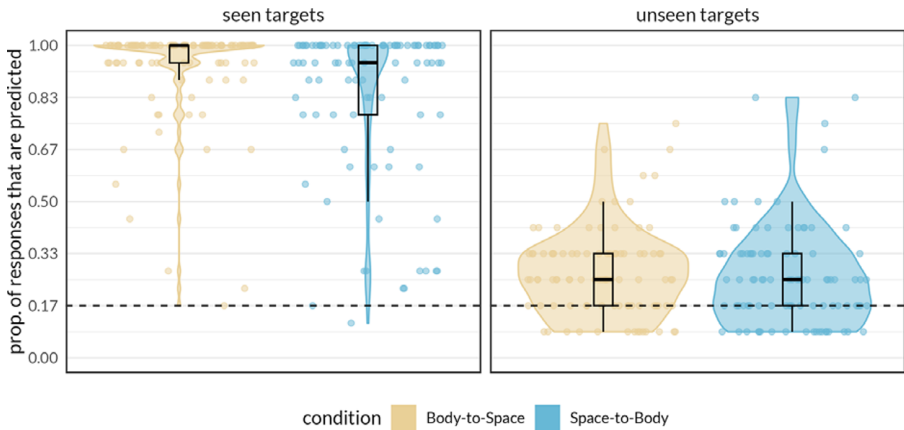
Figure 7. Proportion of predicted responses in the communication phase of Experiment 2. The predicted response for a seen target is simply the word the participant learned for that target in the learning phase; for unseen targets, the predicted response is the paired concept derived from the World Lexicon of Grammaticalization, see main text. We again found no significant difference between the conditions.

significant interaction between condition and seen ($\beta = 0.61$, SE $= 0.24$, $z = 2.51$, $p = 0.012$), again suggesting that prepositions were harder to learn (recall that a 'predicted' response for seen targets means that the sender responded with the word that they had learned for that target meaning).

3.2.3. Entropy of responses

Figure 8 shows the entropy of sender responses for seen and unseen body parts and prepositions. We analyzed these results using a linear model. The response variable was entropy, predicted by target type (body part or preposition, treatment coded) and seen (treatment coded), with no random effects included because they resulted in a
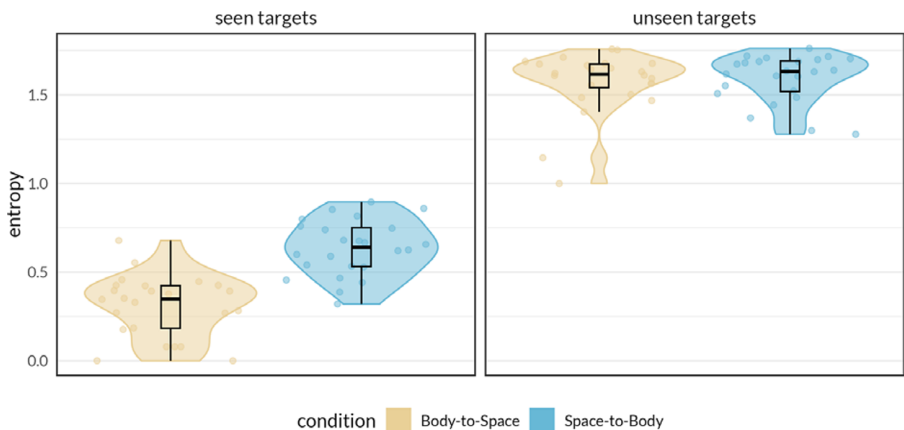

Figure 8. The entropy of sender responses for each target meaning, split by whether the target was seen or unseen. Each dot represents the entropy for one target meaning. Unseen targets had higher entropy than seen targets, and seen prepositions (SPACE-TO-BODY) had higher entropy than seen body parts (BODY-TO-SPACE).

singular fit. As expected, seen targets had lower entropy than unseen ones ($\beta = -1.28$, SE $= 0.047$, $t = -27.54$, $p < 0.001$), but there was no main effect of target type ($\beta = -0.024$, SE $= 0.047$, $t = -0.519$, $p = 0.605$). There was a significant interaction between target type and seen ($\beta = 0.35$, SE $= 0.066$, $t = 5.32$, $p < 0.001$), which again indicates that prepositions were harder to learn, with participants in the SPACE-TO-BODY condition having more variable word choice for trained meanings.

3.2.4. Numeric analysis of association frequencies

Figure 9 shows heatmaps of the response probabilities for senders and receivers for each of the four lists, showing how frequently participants exploited particular extensions during production and comprehension. If participants were behaving
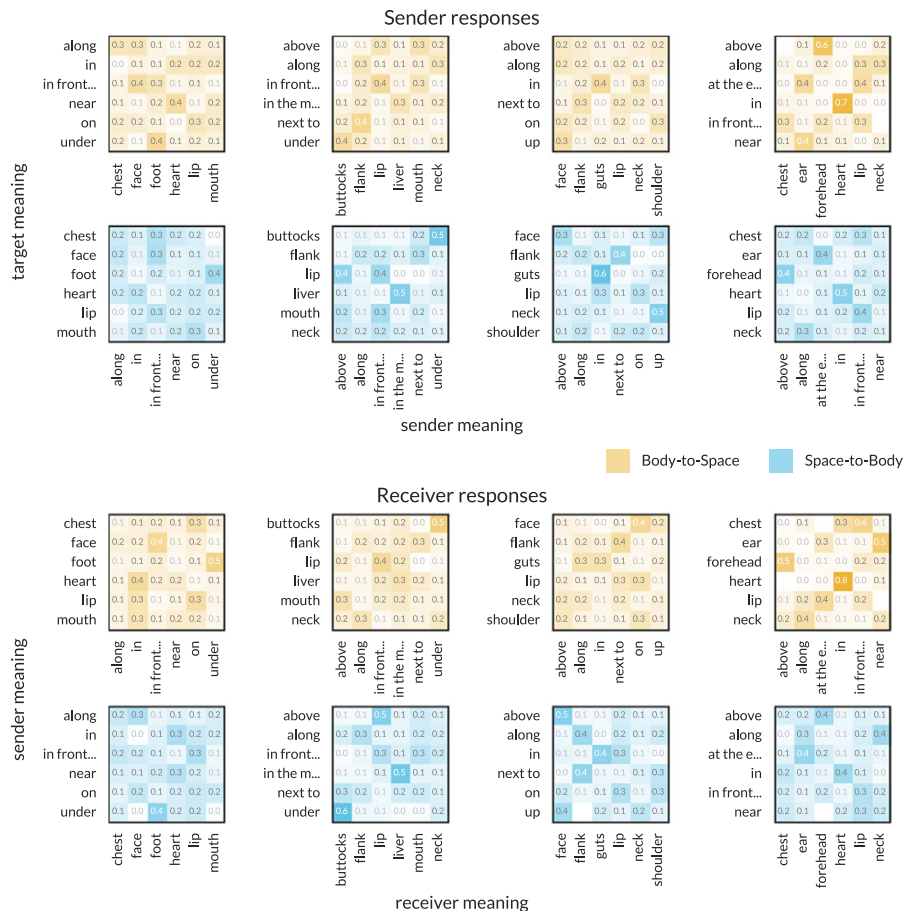
Figure 9. Heatmaps showing response probabilities in unseen trials for both senders and receivers, with colour indicating condition. The sender heatmaps have target meanings on the y-axis and sender responses on the x-axis (indicated by the meaning the participant originally learned for the word they responded with). The receiver heatmaps have the meaning of the word the sender said on the y-axis, and the receiver's response on the x-axis. The values represent the probability of responding with meaning m_1 given the observed meaning m_0, so each row in each subplot sums to 1.

as predicted, we would expect both senders and receivers to show a clear pattern of extension in the BODY-TO-SPACE (yellow) heatmaps, with one or two darkly shaded squares in each row. In the SPACE-TO-BODY (blue) heatmaps, we would expect greater uncertainty, with relatively light squares in every row. Instead, the certainty in both conditions is fairly equally mixed, showing no asymmetry. Some of the expected associations are present in the participants' data, for example, there is a relatively strong association between 'buttocks' and 'under'. But this association is symmetric, i.e., participants are as likely to interpret the word for 'buttocks' as meaning 'under' as the reverse. Some expected associations are present and asymmetric in the expected direction (e.g., senders in list 2 are more likely to say 'flank' when they need to refer to 'next to' than they are to say 'next to' when they need to refer to 'flank'), but other expected associations are much less clear (e.g., senders in list 4 were almost equally likely to say 'ear' for 'at the edge of', the expected extension and 'near'), or asymmetric in the 'wrong' direction (e.g., senders in list 2 were more likely to say 'in the middle of' to refer to 'liver' than vice versa).

3.3. Discussion

As in Experiment 1, we found no evidence of an asymmetry in how individuals perform semantic extension between lexical concepts (body parts) and grammatical ones (spatial prepositions). Participants did not appear to find one direction of semantic extension to be easier or more intuitive than the other.

Many of the participants in Experiment 2 found the task very challenging. This is reflected in their low communicative success for unseen targets (recall Figure 6), as well as in the comments they submitted. One participant said, 'it was quite confusing when all of a sudden the words had different meanings with no explanation'. Another seemed to think the introduction of unseen meanings was done in error: 'I did encountered [*sic*] a difficulty along the way – around half way through phase two it would give a word such as viku and then for the translation have the options of 'under behind beside etc' It must have happened around 10–12 times in the second half of phase 2'. We did not anticipate that the task would be so confusing, and this may have affected our results. However, there were participants who understood the task and had relatively high communicative success, and the lack of asymmetry remains even when considering only their results. We performed the same analyses as before, excluding any participants whose correctness on unseen trials was lower than the median correctness for unseen trials (0.25) and excluding their partners. This gave us results for 94 participants, 46 in the BODY-TO-SPACE condition and 48 in the SPACE-TO-BODY condition. Again, there was no main effect of condition on receiver correctness ($\beta = -0.12$, SE $= 0.16$, $z = -0.78$, $p = 0.44$) or sender's use of predicted responses ($\beta = -0.094$, SE $= 0.24$, $z = -0.40$, $p = 0.69$), and no main effect of target type on sender response entropy ($\beta = -0.018$, SE $= 0.072$, $t = -0.25$, $p = 0.803$).

The combined results of our two experiments suggest that the unidirectional pattern of change from body part terms to spatial prepositions might not be due to a strong asymmetry in how individuals associate the two concepts, since we find no evidence for such asymmetries of association in either of our experiments.

If an asymmetric association between body parts and spatial relations is not responsible for unidirectionality, what else could explain it? In both of our experiments, while we had slightly more body part meanings than prepositions in our

stimulus set, we ensured that body parts and spatial prepositions appeared with equal numbers for each pair (i.e., 6 body parts, 6 prepositions). It could be that this is not an accurate representation of the conceptual space. For example, if natural languages tend to have more labels for body parts than there are meanings to represent in the domain of spatial prepositions, and fewer labels for spatial prepositions than there are concepts for body parts, intuitively this would make extension from body part to spatial preposition easier: a speaker could choose a body part to uniquely identify a spatial preposition, but there would be insufficient unique spatial prepositions to allow them to unambiguously convey body part concepts using spatial prepositions. In the following section, we use a computational model to see how our results might differ if the sizes of the two categories are unequal.

4. Modeling the sender–receiver task

To see whether the size of the category, rather than an asymmetry in associations, could be responsible for asymmetric semantic extension, we designed a model to simulate a modified version of Experiment 2. The model we use is based on the Rational Speech Act (RSA) framework (Frank & Goodman, 2012; Goodman & Frank, 2016). RSA simulates pragmatic communication between two agents, and has been used to study implicature (Bergen et al., 2016), metaphor understanding (Kao et al., 2014), politeness (Yoon et al., 2020), and many other phenomena (see Degen, 2023 for a thorough review). The pragmatic speaker chooses what to say by reasoning about their internal model of a literal listener, and the pragmatic listener agent infers the speaker's intended meaning by reasoning about their internal model of a pragmatic speaker. Reasoning about the interlocutor is achieved via Bayesian inference.

4.1. Model

We present an extension of RSA that allows us to model the sender–receiver task from Experiment 2. We first formulated an extension to the RSA model to allow us to model similarity-based extension of existing terms, then we fit the parameters to our experimental data, meaning that agents in our model have the same (symmmetric) extension preferences as our experimental participants – the details are in Supplementary Materials. All the code related to the model is available on OSF.[7]

We modify the vanilla RSA model to enable communication about a meaning neither the speaker nor listener have a word for in their lexicon, by incorporating a matrix of similarities with an entry for each meaning–meaning pair. The value of $sim(m_1, m_2)$ indicates how strongly associated m_1 is to m_2. Asymmetric associations can be represented by $sim(m_1, m_2) \neq sim(m_2, m_1)$. When an agent has to communicate about a meaning m that they do not have a word for, they rely on the similarities between m and the other meanings. These similarities are parameters in our model, which will be derived from human data.

The lexicon $\mathcal{L}$ is a matrix with one entry for every word–meaning pair. $\mathcal{L}(w, m)$ is 1 if word w has the meaning m and 0 otherwise. The parameter ε determines the

---

[7]https://osf.io/s8hc4.

Table 5. Parameters and functions used in the model of the sender–receiver task from Experiment 2

Symbol	Description
$\mathcal{L}$	Shared lexicon
sim	Similarity matrix
ε	Noise parameter
L_0	Literal listener
S_1	Pragmatic speaker (sender)
L_1	Pragmatic listener (receiver)

amount of noise in the literal listener's distribution over intended meanings given a word (i.e., lexical errors). The components of the model are shown in Table 5.

To match the experimental setup, both agents have the same lexicon: one word for each body part and no spatial prepositions in the BODY-TO-SPACE condition, and vice versa in the SPACE-TO-BODY condition.

While the model described here can be used for both seen and unseen trials, here we only discuss unseen trials. Seen trials are trivial for the model as it has perfect memory, and they provide no insight into semantic extension.

4.1.1. Literal listener

The literal listener observes a word w and outputs a conditional probability distribution over all meanings in the context c, given w, c, and their lexicon $\mathcal{L}$. The literal listener L_0 assigns a probability to some meaning m according to Equation 1. Following the design from Experiment 2, the context is either all body parts or all spatial prepositions.

$$P_{L_0}(m|w,c,\mathcal{L}) \propto \begin{cases} 1-\varepsilon & \text{if } m \text{ in } c \text{ and } \mathcal{L}(w,m)=1 \\ sim(m',m)-\varepsilon & \text{if } m \text{ in } c \text{ and } \mathcal{L}(w,m')=1 \text{ and } sim(m',m)>0 \\ \varepsilon & \text{if } m \text{ in } c \\ 0 & \text{otherwise} \end{cases}$$

(1)

4.1.2. Pragmatic speaker (sender)

The pragmatic speaker S_1 selects a word by performing Bayesian inference on their model of the literal listener. We assume that the prior over words is uniform, so only the likelihoods are needed. The probability the speaker assigns to some word w given the target meaning m_T, context c, and lexicon $\mathcal{L}$ is shown in Equation 2.

$$w_S \sim P_{S_1}(w|m_T,c,\mathcal{L}) = \frac{P_{L_0}(m_T|w,c,\mathcal{L})}{\sum_{w'} P_{L_0}(m_T|w',c,\mathcal{L})}$$

(2)

4.1.3. Pragmatic listener (receiver)

The pragmatic listener performs Bayesian inference on their model of the speaker to compute a distribution over meanings in the context. Again, only the likelihoods are needed since we assume a uniform prior over meanings. The inferred meaning m_R is chosen by sampling from the distribution over all meanings m in the context c given by P_{L_1}, given the word the sender chose w_S. The formula is shown in Equation 3.

$$m_R \sim P_{L_1}(m|w_S, c, \mathcal{L}) = \frac{P_{S_1}(w_S|m, c, \mathcal{L})}{\sum_{m' \in c} P_{S_1}(w_S|m', c, \mathcal{L})} \quad (3)$$

4.2. Fitting the model

We fit this model to our experimental data from Experiment 2, predicting sender behaviour (which label was selected to convey a given target meaning by the pragmatic speaker) and receiver behaviour (which meaning was selected given a particular signal by the pragmatic listener). Since the training lexicon is fixed, the parameters we fit to experimental data are the underlying similarity matrix, and the noise parameter epsilon – i.e., we infer from the experimental data what similarity matrix and what noise parameter would account for our experimental data, under the assumption that our participants are behaving in the same way as our RSA speaker and listener. Further details of the fitting procedure are given in Supporting Materials.

4.3. Manipulating the sizes of the categories

We used the fitted model to run a simulation of a modified version of Experiment 2 using the same four lists the human participants saw, but manipulated how many body parts or prepositions there were – either six body parts and four prepositions or four body parts and six prepositions. To generate the lists of four meanings, we simply picked four randomly each time from the list of six. This was done for both directions of extension, BODY-TO-SPACE and SPACE-TO-BODY. The resulting proportions of receiver success from the model are shown in Figure 10.

Correctness was above chance in both directions of extension, because agents are able to exploit the same meaning–meaning associations as our human participants. Communication is more successful when there are more words than meanings
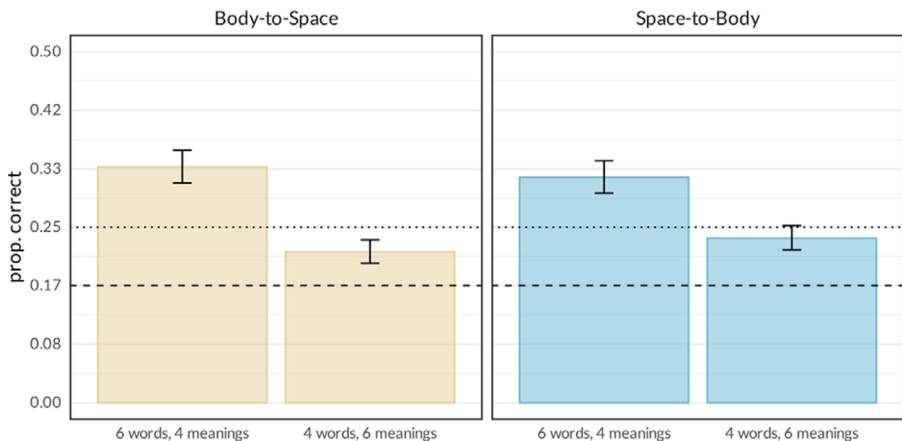
Figure 10. Proportion of trials where the model's receiver produced the correct response. Responses were generated 100 times for each target meaning, each list, each direction (BODY-TO-SPACE and SPACE-TO-BODY), and each manipulation. The error bar shows the 95% confidence interval. The dotted line shows chance performance, which varies depending on the number of meanings the receiver selects from. Receiver success was higher when agents had more words and fewer meanings to refer to than when they had fewer words and more meanings. Direction had no effect on success.

because it is possible to uniquely assign each label to one meaning, whereas when there are fewer words than meanings, at least one label must be used to refer to multiple meanings and is therefore ambiguous. This effect was seen both in BODY-TO-SPACE and SPACE-TO-BODY extension – in other words, moving to unequally sized lexicons does not reveal some inherent asymmetry in the similarity matrices, but does confirm (using well-motivated similarities) that it is easier to extend more terms to cover fewer concepts than the reverse.

4.4. Discussion

Simulated agents found it easier to successfully extend existing words when they had more words available to label fewer concepts. However, the direction of extension (BODY-TO-SPACE and SPACE-TO-BODY) had no effect, confirming our finding that there is no evidence in our experiment that participants have asymmetric associations between body parts and spatial prepositions.

   This result suggests that unidirectionality could be due to asymmetric semantic extension caused by differences in the size of the sets of lexical terms and grammatical terms, rather than requiring asymmetries in associations between the two domains. In particular, this would require that languages tend to have more lexical terms than grammatical terms in the domains in which unidirectional grammaticalization typically occurs (e.g., in our case, more body part terms than spatial prepositions). This is plausible, given that lexical classes are open and new words are often added to them, while grammatical classes are less likely to be added to.

5. General discussion

Using two behavioural experiments, we tested whether individuals have a unidirectional preference for using lexical items to refer to grammatical concepts when performing semantic extension. We found no evidence of this preference. Thus, we have no evidence supporting the widely-held assumption that such asymmetries in association are the cause of the observed unidirectional tendency of grammatical change, contrary to the assumptions of CMT informed explanations of unidirectionality (Lakoff & Johnson, 1980), the asymmetric priming hypothesis (Jäger & Rosenbach, 2008), IITSC (Traugott, 1989), and Haspelmath's (1999) explanation of grammaticalization based on the invisible-hand theory of language change (Keller, 1989).

   Given the lack of evidence for an asymmetric bias, we hypothesized that asymmetry in semantic extension could simply be caused by having more lexical items than grammatical items. Using an RSA model of Experiment 2, we showed that having more body part terms than spatial prepositions would lead to asymmetric semantic extension between the two domains. If evidence were found for there being more lexical items than grammatical items cross-linguistically, this would be a plausible explanation for unidirectionality of grammatical change.

   Below, we discuss some limitations of our work and propose extensions that could bring us closer to understanding the cognitive and communicative origins of unidirectionality.

5.1. Limitations and extensions

We only consider semantic extension between body parts and spatial prepositions. Given the strong evidence for asymmetric conceptual associations between space and time (Boroditsky, 2000; Bottini & Casasanto, 2013), it seems plausible that the historical unidirectional extension of spatial terms to temporal concepts is because speakers perform semantic extension asymmetrically between the two domains. Our methods could be used to test for this asymmetry using space and time stimuli.

Words in our experiments were used in isolation, not in a sentence or phrases. This may have made the semantic extension task more difficult because participants could not rely on linguistic context to determine a word's intended meaning. In the Invited Inferencing Theory of Semantic Change (Traugott & Dasher, 2001), context is thought to be the key that allows interlocutors to understand the new meaning of a word through pragmatic inference. Adding context may reduce the difficulty of the task, potentially revealing asymmetries our current method obscures.

We require participants to extend from body parts to spatial prepositions without any intermediate stages, but it is unlikely that the semantic change component of grammaticalization really happens this suddenly. Heine (1997, p. 44) suggested four stages in the process of body part to spatial concept extension:

1. Stage 1 – A region of the human body
2. Stage 2 – A region of an (inanimate) object
3. Stage 3 – A region in contact with an object
4. Stage 4 – A region detached from the object

One reason we chose body parts and spatial prepositions is that we thought participants' existing associations between body parts and spatial regions could serve in place of gradualness. To test if explicit gradualness leads to asymmetry in Experiment 2, a third category of meanings could be added between body parts and spatial prepositions (e.g., regions of an object) that participants in both BODY-TO-SPACE and SPACE-TO-BODY would have to communicate about. Note that Bowerman and Smith (2022) successfully had participants chain semantic extensions in an artificial communicative task similar to ours.

Some of the simplifications in our design may have masked inherent asymmetries between domains. As explained in Section 3.1.2, by design, the stimuli lists for Experiment 2 contained no body parts that mapped to similar prepositions. It may be that the associations between these concepts are many-to-one, meaning that many body parts map to one spatial relation while only one spatial relation maps to many body parts. This would make it easier to know which spatial relation someone is referring to when using a body part term than the reverse, leading to asymmetric semantic extension. Future work could investigate this by eliciting human associations between body parts and spatial relations and seeing if they exhibit many-to-one mappings.

The model we presented relies on a similarity matrix derived from our experimental data. It may be more informative to instead use human association ratings for each pair of meanings and use these as parameters to the model. Existing datasets of human similarity/association judgements (e.g., Finkelstein et al., 2002) do not have enough data for associations between body parts and spatial relations to be useful for our research, so new association data would have to be collected.

6. Conclusion

Using experimental methods, we tested the assumption underlying many theoretical explanations of unidirectionality of grammatical change that individuals perform semantic extension asymmetrically between lexical and grammatical domains, thus leading to a historical unidirectional tendency. Unexpectedly, we found no evidence of such an asymmetry. While previous experimental work testing the asymmetric priming hypothesis also ended with a null result, our finding was still unexpected. Future experiments could investigate whether a task with additional linguistic context or gradualness leads to asymmetric semantic extension.

We used a computational model of communication to show that an alternative source of asymmetry would arise if there are more lexical items than grammatical items in the lexicon, which straightforwardly results in a preference to extend labels from the larger lexicon to cover the concepts from the smaller domain.

Our work shows that experimental methods can be an important tool in grammaticalization research, and research on semantic change in general, allowing us to test crucial assumptions underpinning theoretical accounts. These methods allow us to go beyond the corpus data, providing new insight into how individuals shape language, in our case, questioning central tenets of prior work.

References

Bergen, L., Levy, R., & Goodman, N. (2016). Pragmatic reasoning through semantic inference. *Semantics and Pragmatics*, 9, 1–91. https://doi.org/10.3765/sp.9.20

Boroditsky, L. (2000). Metaphoric structuring: Understanding time through spatial metaphors. *Cognition*, 75(1), 1–28. https://doi.org/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, 1–9.

Bowerman, J., & Smith, K. (2022). An experimental study of semantic extension in a novel communication system. In *Proceedings of the 44th annual conference of the cognitive science society* (pp. 163–169). eScholarship University of California.

Brooks, M. E., Kristensen, K., van Benthem, K. J., Magnusson, A., Berg, C. W., Nielsen, A., … Bolker, B. M. (2017). glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *The R Journal*, 9(2), 378–400. https://doi.org/10.32614/RJ-2017-066

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

Culbertson, J., Smolensky, P., & Legendre, G. (2012). Learning biases predict a word order universal. *Cognition*, 122(3), 306–329. https://doi.org/10.1016/j.cognition.2011.10.017

de Leeuw, J. R., Gilbert, R. A., & Luchterhandt, B. (2023). jsPsych: Enabling an open-source collaborative ecosystem of behavioral experiments. *Journal of Open Source Software*, 8(85), 5351. https://doi.org/10.21105/joss.05351

Degen, J. (2023). The rational speech act framework. *Annual Review of Linguistics*, 9(1), 519–540. https://doi.org/10.1146/annurev-linguistics-031220-010811

Finkelstein, L., Gabrilovich, E., Matias, Y., Rivlin, E., Solan, Z., Wolfman, G., & Ruppin, E. (2002). Placing search in context: The concept revisited. *ACM Transactions on Information Systems*, 20(1), 116–131.

Frank, M. C., & Goodman, N. D. (2012). Predicting pragmatic reasoning in language games. *Science*, 336(6084), 998. https://doi.org/10.1126/science.1218633

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, 1–12. https://doi.org/10.3389/fcomm.2023.1136338

Goodman, N. D., & Frank, M. C. (2016). Pragmatic language interpretation as probabilistic inference. *Trends in Cognitive Sciences*, 20(11), 818–829. https://doi.org/10.1016/j.tics.2016.08.005

Haspelmath, M. (1999). Why is grammaticalization irreversible? *Linguistics*, 37(6), 1043–1068. https://doi.org/10.1515/ling.37.6.1043

Heine, B. (1997). *Cognitive foundations of grammar*. Oxford University Press.

Heine, B. (2003). Grammaticalization. In B. Joseph & R. Janda (Eds.), *The handbook of historical linguistics* (pp. 575–601). John Wiley & Sons, Incorporated.

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

Heine, B., & Kuteva, T. (2012). Grammaticalization theory as a tool for reconstructing language evolution. In K. R. Gibson & M. Tallerman (Eds.), *The Oxford Handbook of Language Evolution* (pp. 512–527). Oxford Academic. https://doi.org/10.1093/oxfordhb/9780199541119.013.0054

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

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

Hollenbach, B. E. (1995). Semantic and syntactic extensions of body-part terms in mixtecan: The case of 'face' and 'foot'. *International journal of American Linguistics*, 61(2), 168–190. https://doi.org/10.1086/466250

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

Issah, S. A. (2020). On body-part nouns and the expression of spatial and locative reference in dagbani. *Topics in Linguistics*, 21(2), 41–61. https://doi.org/10.2478/topling-2020-0009

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

Kanwal, J., Smith, K., Culbertson, J., & Kirby, S. (2017). Zipf's law of abbreviation and the principle of least effort: Language users optimise a miniature lexicon for efficient communication. *Cognition*, 165, 45–52. https://doi.org/10.1016/j.cognition.2017.05.001

Kao, J., Bergen, L. & Goodman, N. (2014). Formalizing the pragmatics of metaphor understanding. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 36(36), 719–724.

Karjus, A., Blythe, R. A., Kirby, S., Wang, T., & Smith, K. (2021). Conceptual similarity and communicative need shape colexification: An experimental study. *Cognitive Science*, 45(9), e13035. https://doi.org/10.1111/cogs.13035

Keller, R. (1989). Invisible-hand theory and language evolution. *Lingua*, 77(2), 113–127. https://doi.org/10.1016/0024-3841(89)90011-9

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

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmertest package: Tests in linear mixed effects models. *Journal of Statistical Software*, 82(13), 1–26. https://doi.org/10.18637/jss.v082.i13

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

Meillet, A. (1912). L'évolution des formes grammaticales. *Scientia (Rivista di Scienza)*, 12(6), 384–400.

Norde, M. (2009). *Degrammaticalization*. Oxford University Press.

R Core Team. (2023). *R: A language and environment for statistical computing* [Computer software manual]. Vienna, Austria. Retrieved from https://www.R-project.org/

Saldana, C., Oseki, Y., & Culbertson, J. (2021). Cross-linguistic patterns of morpheme order reflect cognitive biases: An experimental study of case and number morphology. *Journal of Memory and Language*, 118, 104204. https://doi.org/10.1016/j.jml.2020.104204

Silvey, C., Kirby, S., & Smith, K. (2015). Word meanings evolve to selectively preserve distinctions on salient dimensions. *Cognitive Science*, 39(1), 212–226. https://doi.org/10.1111/cogs.12150

Smithson, M., & Verkuilen, J. (2006,March). A better lemon squeezer? Maximum-likelihood regression with beta-distributed dependent variables. *Psychological Methods*, 11, 54–71. https://doi.org/10.1037/1082-989X.11.1.54

Suttle, L., & Goldberg, A. E. (2011). The partial productivity of constructions as induction. *Linguistics*, 49(6), 1237–1269. https://doi.org/10.1515/ling.2011.035

Svorou, S. (1994). *Grammar of space*. John Benjamins Publishing Company.

Sweetser, E. E. (1988). Grammaticalization and semantic bleaching. *Annual meeting of the Berkeley Linguistics Society*, 14, 389–405.

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

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

Traugott, E. C., & Trousdale, G. (2010). Gradience, gradualness and grammaticalization: How do they intersect? In E. C. Traugott & G. Trousdale (Eds.), *Gradience, gradualness and grammaticalization* (pp. 19–44). John Benjamins Publishing Company.

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). Cognitive Science Society.

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

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

Yoon, E. J., Tessler, M. H., Goodman, N. D., & Frank, M. C. (2020). Polite speech emerges from competing social goals. *Open Mind*, 4, 71–87. https://doi.org/10.1162/opmi_a_00035