

UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

An efficiency-based effect of frequency on lexicalization: a dyadic experiment

Permalink

<https://escholarship.org/uc/item/1h4173zd>

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 48(0)

Authors

Uegaki, Wataru

Mucha, Anne

Qing, Ciyang

Publication Date

2026

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

An efficiency-based effect of frequency on lexicalization: a dyadic experiment

Wataru Uegaki (w.uegaki@ed.ac.uk), Anne Mucha & Ciyang Qing

School of Philosophy, Psychology and Language Sciences,
The University of Edinburgh

Abstract

Zipf's Law of Abbreviation states that the more frequent a word is, the shorter it tends to be. Zipf's own explanation of the law in terms of the trade-off between cognitive effort and communicative accuracy, i.e., Principle of Least Effort, predicts that a similar pattern holds not only for word length but also for the lexical/compositional distinction. That is, lexical forms tend to express more frequent meanings compared to compositional forms. In this paper, we report on a dyadic communication experiment that supports this prediction.

Keywords: Zipf's law of abbreviation; Principle of Least Effort; Compositionality; Lexicalization; Dyadic experiments; Linguistic convention formation

Introduction

A number of cross-linguistic generalizations have been accounted for by the idea that languages are optimized to support *efficient communication* (Aylett and Turk 2004; Denić et al. 2022; Ferrer-I-Cancho and Solé 2003; Imel et al. 2026; Kemp et al. 2018; Kirby et al. 2015; Levshina 2022, among many others). A prominent example of such a generalization is Zipf's Law of Abbreviation (ZLA), according to which the more frequent a word is, the shorter it tends to be (Zipf, 1936). Zipf offered an efficiency-based explanation for this pattern, Principle of Least Effort (PLE), formulated based on the trade-off between cognitive effort and communicative accuracy (Zipf, 1949). Using shorter words saves effort but sacrifices accuracy (the space of meanings one can express with short forms is limited; short forms are vulnerable to noise); Using longer words increases accuracy, but incurs more effort. An optimal solution for this dilemma is to allocate shorter forms to frequent meanings and longer forms to rare ones, so the average token length will be shorter. As detailed below, PLE yields a new hypothesis about the effect of frequency on lexicalization and compositionality: roughly, the combined effect of cognitive effort and communicative accuracy yields a pattern where *lexical forms tend to express more frequent meanings compared to compositional forms*. A potential manifestation of this effect is the cross-linguistic lexicalization pattern in expressions for impossibility (Mucha et al., 2024; Uegaki et al., 2025). When a language has a lexical modal auxiliary for impossibility, the item tends to express a more frequent, *non-epistemic*, flavor of impossibility, rather than a less frequent, *epistemic*, flavor. Epistemic impossibility tends to be expressed by a compositional form, e.g., a combination

of negation and a possibility modal. (See the Hypothesis section for details of the cross-linguistic pattern.)

Here, we report on a dyadic communication experiment (Kanwal et al., 2017) that supports our hypothesis on the effect of frequency on lexicalization. In our experiment, participants were given a language in which there are two candidate forms—one lexical and one compositional—that can be mapped to a high frequency meaning and a low frequency meaning. We tested the participants' choice of forms in conveying the meanings, under four conditions that vary in (a) the presence/absence of interaction and (b) the presence/absence of additional cognitive effort for using the compositional form (as opposed to a lexical form). Our results confirm that both interaction and the cognitive effort drive the participants' choice to lexicalize the high frequency meaning. Our study offers a new experimental support for the role of communicative efficiency in lexicalization, along with existing analyses of cross-linguistic patterns (e.g. Khishigsuren et al., 2025; Xu et al., 2024). Our study has further implications for the understanding of cognitive effort underlying PLE. Existing dyadic experiments on PLE use time pressure to operationalize effort (Kanwal et al., 2017). Our results show that formal properties of stimuli in the absence of time pressure can drive effort.

Background

Zipf (1936) famously observed that the length of a frequent word tends to be shorter. To explain this inverse correlation between word length and frequency, Zipf (1949) postulated the Principle of Least Effort, according to which an effective communicative system should generally favor forms that require less effort to produce, so long as the intended meanings can be accurately communicated.

Kanwal et al. (2017) used a dyadic Artificial Language Learning experiment to directly test the explanation of ZLA in terms of PLE. In the training phase, participants were presented with two distinct objects (with one being more frequent than the other) and trained on their respective long names and shared short name. In the testing phase of the condition that combines pressures from both accuracy and effort, participants were paired with a partner to play a communication game. In each trial, the Sender saw an object that they needed to describe to their partner (the Receiver) by choosing either its long name or the short name. The time it took to transmit the chosen name is proportional to its length. When the Receiver received the message, they needed to choose which

object was intended by the Sender. Participants were told that they would receive a bonus if they completed the game the most quickly with the highest accuracy, and therefore both accuracy and length mattered for the task. In the other three conditions, one or both of the pressures were removed. In conditions that removed the pressure for accuracy, there was no communication and participants were simply tasked to choose a name to describe the object. In conditions that removed the pressure for less effort, there was no time difference between long and short names.

Kanwal et al. found that participants preferentially assigned the more frequent meaning to the short form only when both communicative accuracy and production effort were jointly relevant. This pattern provides direct experimental support for an efficiency-based explanation of ZLA grounded in PLE. At the same time, word length represents only one possible source of effort. If PLE is understood more broadly as a constraint on cognitive effort in communication, it should extend beyond word length to other dimensions of linguistic complexity. In this paper, we pursue this more general interpretation of PLE by testing a prediction about competition between lexical and compositional forms.

Hypothesis

Our hypothesis can be stated as follows:

- (1) If a language contains a compositional form F_c and a lexical form F_l as candidate forms for a frequent meaning M_{freq} and an infrequent meaning M_{inf} , the combined effect of accuracy and effort results in the mapping of F_l to M_{freq} and of F_c to M_{inf} .

The hypothesis follows from PLE, assuming that using a compositional form incurs additional cognitive effort relative to using a lexical form (due to e.g., the need to retrieve multiple lexical items from memory; the need to derive the complex form through accurate composition of lexical items, including potentially covert scopal configurations). Note that the hypothesis is in principle independent from ZLA as an empirical observation since length is independent from lexicalization—a compositional form may be shorter than a lexical form.

A potential empirical manifestation of this hypothesis is a cross-linguistic lexicalization pattern of impossibility, documented in Mucha et al. (2024) and Uegaki et al. (2025). Some natural languages have lexicalized impossibility modals, in addition to a compositional form that combines negation with a possibility modal. The relevant empirical pattern is that lexicalized impossibility modals tend to express non-epistemic (specifically deontic) modality, more likely than epistemic modality. This is illustrated with examples from Hausa (Afro-Asiatic) and Thai (Kra-Dai) in (2) and (3). Both languages use lexicalized impossibility modals (*kada* in Hausa, *harm* in Thai) to express deontic impossibility (2). As shown in (3), however, in both languages, a compositional form—consisting of morphological negation and a possibil-

ity modal—must be used to convey epistemic impossibility.¹

- (2) *Deontic impossibility*: “Visitors can’t stay after 6pm.”
 - a. **Hausa**: *Kada mazi yarta su wuce karfe 6*
 $\neg \diamond$ visitors 3pl.sbjv stay hour 6
 - b. **Thai**: *Yardpubuay harm yuu lang hok mornng*
 patient.visitor $\neg \diamond$ exist after six o’clock
- (3) *Epistemic impossibility*: “Ben can’t be at home (based on his habit).”
 - a. **Hausa**: *Ba zai yiwu Ben ya kasance a gida ba*
 NEG 3sg.fut $\diamond$ Ben 3sg be at house NEG
 - b. **Thai**: *Ben yuu tii baan mai dai*
 Ben exist LOC house NEG $\diamond$

The apparent asymmetry in the lexicalization of so-called ‘modal flavors’ may at least partly be grounded in a frequency contrast as formulated in (1). Corpus studies suggest that epistemic uses of modals are less frequent than non-epistemic uses (e.g. Van Dooren et al. 2022; Veselinovic 2019), and also that deontic impossibility in particular is expressed more frequently than epistemic impossibility (Bell et al., 2024). Our study tests (1) by observing lexicalization in a communicative setting that involves a frequency bias.

Methods and Predictions

Prior to implementation, the experiment was preregistered on OSF ([link to the preregistration](#), which includes motivations for the design and the statistical analysis).

Design

In a dyadic setting (cf. Kanwal et al. 2017), we investigate the combined effect of accuracy and cognitive effort on lexicalization, given a language where users have a choice of using a lexical and a compositional form. To do this, we test lexicalization in 4 between-subject conditions that vary in two factors: (i) the presence/absence of interaction with a partner [$\pm$ part(ner)]; and (ii) the presence/absence of a distractor expression [$\pm$ dist(ractor)]. The presence of interaction increases the pressure from accuracy. The distractor variable manipulates the cognitive effort for the compositional form relative to the lexical form, as explained below.

Materials and Procedure

The experiment consists of an exposure phase followed by a testing phase. In the exposure phase, participants observe 40

¹ As pointed out by a reviewer, some Thai speakers judge the sentence in (3b) as strictly non-epistemic, and thus incompatible with the intended reading. In line with these judgments, relevant discussions in the literature (Simpson, 2001; Visonyangoon, 2000), describe *dai* as a root modal with deontic and ability readings. However, according to the data provided by our native speaker consultant, *dai* is compatible with epistemic readings when combined with negation (similarly to English *can*). Our general argument is not affected by this variation, since alternative expressions of epistemic impossibility are also compositional.

1. Lexicalized negative existential (nex) sentence, e.g., ***Nothing** in this picture is red.*
 2. Compositional negative existential (nex) sentence, e.g., ***Everything** in this picture is **not** red.*
 3. **In [+dist]:** distractor weak sentence, e.g., ***Not everything** in this picture is red.*
- In [-dist]:** non-distractor weak sentence: *There are multiple objects in this picture.*

The pictures consist of the following four types:

1. *circles* where nex is true (henceforth denoted as $\neg\exists_O$);
2. *triangles* where nex is true ($\neg\exists_\Delta$);
3. *circles* where only the weak sentence is true ($\neg\forall_O$);
4. *triangles* where only the weak sentence is true ($\neg\forall_\Delta$).

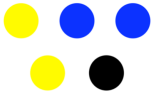
One shape is 3 times more frequent than the other, both in exposure and testing. The frequency of the shape is counter-balanced between circles and triangles. In the exposure phase, $\neg\forall$ pictures are always presented with weak sentences, and $\neg\exists$ pictures are always presented with nex sentences.²

The testing phase contains 40 trials in the [–part] conditions and 80 trials (40 Sender; 40 Receiver tasks) in the [+part] conditions. In the [+part] conditions, participants interact with a partner (see Fig. 1b). They alternate between being the Sender, who chooses a sentence given a picture, and the Receiver, who chooses a picture given the sentence the Sender has chosen. They get feedback on whether the picture the Sender saw is the same as the picture the Receiver chose. In [–part] conditions, participants simply choose a sentence given a picture, and get feedback on whether the sentence they have chosen is accurate according to the exposure. Thus, by design, communicative success in the [+part] conditions requires resolving the ambiguity with respect to shape (i.e., circles vs. triangles), but [–part] tasks can succeed without shape disambiguation. In all conditions, participants receive +1 point if the picture is matched correctly, and –1 point otherwise. In the [+part] conditions, points are awarded by pair, so that paired participants work together towards a total score. After finishing the experiment, bonus payment is given according to the final score.

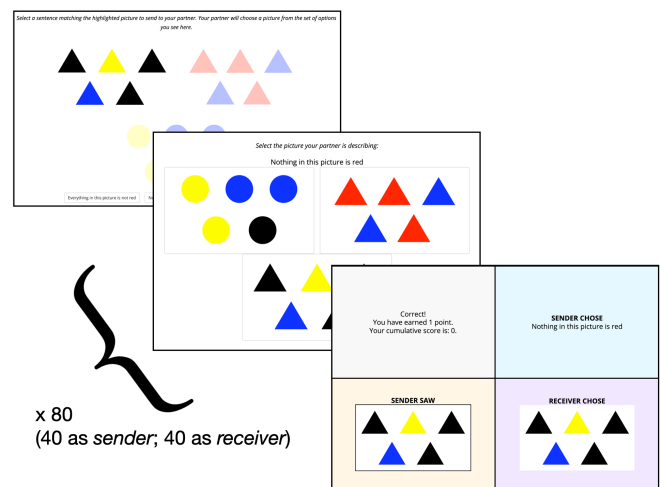
Participants

We collected data from 151 US- and UK-located monolingual English speakers (non-colorblind), recruited on Prolific (38 in the [+part,+dist] condition; 34 in the [+part,-dist] condition;

²Note that the compositional nex form is in principle covertly scope-ambiguous. In addition to the negative existential surface scope reading, it has an inverse scope reading which is semantically equivalent to the weak form in the [+dist] condition, and thus compatible with $\neg\forall$ pictures. However, exposure does not contain picture-sentence pairs that verify inverse scope for compositional nex. This covert scope ambiguity provides additional cognitive effort for the compositional form (on top of other factors), which can then be boosted in the presence of a distractor, as explained below.

$\neg \exists$ pic with lexical nex form  Nothing in this picture is red x 12	$\neg \exists$ pic with compositional nex form  Everything in this picture is not red x 12	$\neg \forall$ pic with weak form  Not everything in this picture is red x 6
 Nothing in this picture is red x 4	 Everything in this picture is not red x 4	 Not everything in this picture is blue x 2

(a) Exposure format and input frequency



(b) Testing format

Figure 1: Format of the experiment in the [+part,+dist] condition. In [-part] conditions, the testing only involves 40 Sender trials. In [-dist] conditions, the weak form is always *There are multiple objects in this picture*. 4 predicates were used: *is red*, *is blue*, *is filled*, *is broken*.

39 in the [-part,+dist] condition; 40 in the [-part;-dist] condition). Participants were paid £1.8 for [-part] and £7 for [+part] conditions (commensurate with the study durations), in addition to the bonus payment.

Distractor and cognitive effort

The presence of the distractor weak form in [+dist] conditions is designed to increase the cognitive effort for choosing the *compositional* nex form, relative to the *lexical* nex form. This is analogous to the time-pressure manipulation in Kanwal et al. (2017), which increased the task effort for choosing the long forms. As suggested above, a factor that contributes to the cognitive effort in using a compositional form, relative to a lexical form, is the need for syntactic and semantic compo-

sition. We take *order-sensitivity* of an item to be indicative of such need—if the semantics of the item depends on the order of its component elements, it indicates that the item is composed of multiple morphemes, rather than being a lexical unit.

In our design, the distractor form involves the same lexical elements as the compositional nex form, but in different syntactic configurations, manifesting in a different linear order of lexical elements. This makes salient the compositional nature of the compositional nex form, as mapping the forms to correct meanings encourages attention to their compositional structures, not merely recognizing them as a collection of words. In contrast, the non-distractor form does not involve the same lexical elements as the compositional nex form. Thus, in the non-distractor conditions, the compositional nature of the compositional nex form is not made as salient as in the distractor conditions.

Crucially, this manipulation isolates the cognitive burden of structural composition from raw production length. Because the length difference between the lexical and compositional forms is held constant across both the [+dist] and [−dist] conditions, any behavioral differences between these groups can be attributed to the compositional complexity made salient by the distractor, rather than the mere length of the expressions.

Predictions

Our hypothesis is that both interaction and distractor contribute to the lexicalization of the higher frequency meaning. Our dependent variable is thus whether the *lexical* nex form is used. The independent variables are (a) whether the target picture contains the frequent shape (IS.FREQUENT) and (b) the CONDITIONS ([+part,+dist], [+part,−dist], [−part,+dist], [−part,−dist]). Only the responses in the Sender trials are analyzed in the [+part] conditions, following Kanwal et al. (2017). In the [−part] conditions, all testing trials are analyzed. The relevant trials are those with $\neg\exists$ pictures, and we remove weak responses from the data analysis.

Given the above, we predict a frequency bias in the [+part,+dist] condition. The pressure for accuracy that comes from interaction with a partner is expected to drive disambiguation of the quantifier domain (i.e., mapping each nex form to either circles or triangles). In accordance with (1), we predict that pairs will converge on the *lexical* nex form for the more frequent shape. We expect that this bias will be reduced if the distractor or the interaction is removed, in particular: (i) In the [−part,+dist] condition, the pressure for disambiguation is removed, since without interaction, the use of any nex form in $\neg\exists$ trials will lead to success, irrespective of the shape. We thus expect a lower lexicalization rate for the frequent shape compared to the [+part,+dist] condition. (ii) In the [+part,−dist] condition, the cognitive effort associated with the compositional nex form is reduced. We predict that this also reduces the bias towards the lexical form for the frequent shape. (iii) We expect the bias reduction in the [−part] conditions to be modulated by the type of weak sentences. The difference between [+part] and [−part] is predicted to be larger in the

presence of a distractor ([+part,+dist] vs. [−part,+dist]) compared to the conditions with non-distractor weak sentences.

In addition to the above predictions, we have preregistered an exploratory hypothesis according to which interaction with a partner has an effect also in the absence of a distractor. That is, we check whether there is an interaction between IS.FREQUENT and [+part,−dist] relative to [−part,−dist].

Results

The results are shown in Fig. 2 and Fig. 3. We fitted a binomial Generalized Linear Mixed Model predicting the binary outcome of lexical form use, with by-participant and by-item random effects.³ We used [+part, +dist] as the baseline for Condition so that the first two predicted effects would directly correspond to the model's coefficients (see Table 1). The OSF stores [the stimuli, data, and analysis code](#).

All three predicted effects discussed above are significant: (i) When the distractor weak form was present, participants without a partner were less likely to use the lexical form for the more frequent object than those with a partner ($\beta = -1.697, p < .001$); (ii) Among the participants with a partner, when the distractor form was absent, participants were less likely to use the lexical form than when the distractor form was present ($\beta = -0.877, p < .001$); (iii) The effect of having a partner on the frequency bias for the lexical form is modulated by the presence or absence of a distractor form: when the distractor was present, the difference between participants with and without a partner in terms of their preference to use the lexical form for the frequent object is larger than when the distractor was absent ($\beta = 1.19, p < .001$).⁴ As for the exploratory analysis, we found that even when the distractor was absent, participants with a partner were more likely to use the lexical form for the more frequent object than those without a partner ($\beta = 0.502, p = 0.036$).

Discussion

The results confirm our hypothesis in (1), providing new support for PLE as a general principle governing optimization of semantic systems. Our study has further implications for the understanding of cognitive effort underlying PLE. Existing dyadic experiments on PLE typically use time pressure—which is inherently tied to word length—to operationalize effort (Kanwal et al., 2017). In contrast, our results demonstrate that the formal, structural properties of a language can independently drive effort, even in the absence of explicit time pressure. Importantly, the mere length difference between the lexical and compositional forms cannot explain our results, as

³Our model, following Kanwal et al.'s (2017), includes Condition as a 4-level variable as a fixed-effect term. One may equivalently (modulo the random-effect structure) include Partner and Distractor as 2 binary variables and include full interactions among object frequency, Partner and Distractor in the fixed effects.

⁴This coefficient is derived from a combination of the model's interaction terms, calculating the difference between the partner effect in the distractor condition, $\beta = -1.697$, and the partner effect in the non-distractor condition, $\beta = -0.877 - (-1.379) = 0.502$.

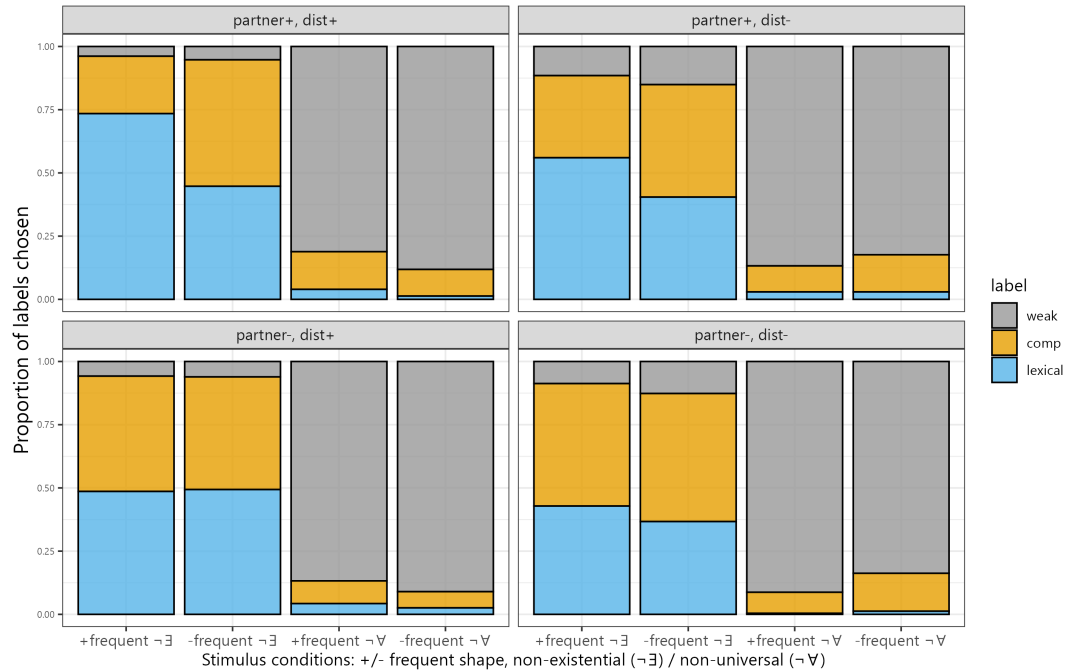


Figure 2: The proportion of Sender label choices with respect to the 4 picture types in all conditions.

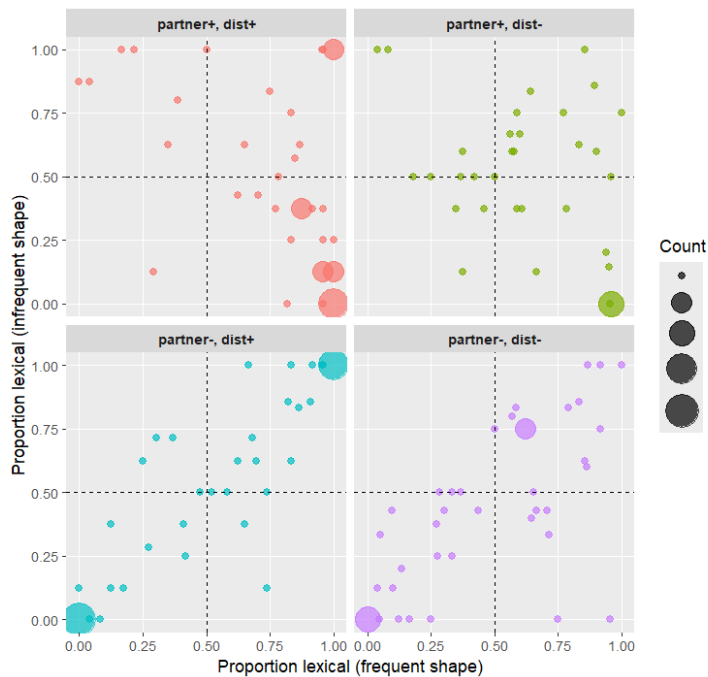


Figure 3: The proportion of trials in which the lexical form was used to label the frequent shape versus the proportion of trials in which the lexical form was used to label the infrequent shape, in all four conditions. Data points represent a participant's director productions of nex labels for $\neg\exists$ pictures. Data points in the bottom right quadrant indicate participants who mostly use the lexical form for the frequent shape and the compositional form for the infrequent shape, data points in the top left quadrant indicate the opposite direction of convergence. Data points in the top right and bottom left corner indicate participants who mostly use the lexical or compositional form, respectively, for both shapes.

the length asymmetry was present equally in both the [+part, +dist] and [+part, -dist] conditions. Instead, it is the presence of a formally similar distractor that increases the processing burden, providing a new dimension of cognitive effort. This is further evidenced by the error rate we observe in the [+part] conditions. Table 2 summarizes the relative frequency of 'force errors' made by Receivers in $\neg\exists$ trials. 'Force error' means that the Sender labels a $\neg\exists$ picture with a

compositional or lexical nex form, but the Receiver responds incorrectly by choosing the $\neg\forall$ picture. In the [+part, +dist] condition, this type of error occurs in almost 15% of trials in which the Sender chose the compositional nex form, whereas the error rates are below 1% in the condition without a distractor, or when the Sender sends the lexicalized nex sentence. This suggests that the compositional form is indeed associated with greater difficulty in form-meaning-mapping in the

	β	SE	p
(Intercept)	-0.153	.299	.608
Obj=freq	1.645	.164	<.001
Cond=[-part, +dist]	0.246	.429	.566
Cond=[+part, -dist]	0.082	.435	.851
Cond=[-part, -dist]	-0.361	.423	.394
Obj=freq&Cond=[-part, +dist]	-1.697	.242	<.001
Obj=freq&Cond=[+part, -dist]	-0.877	.236	<.001
Obj=freq&Cond=[-part, -dist]	-1.379	.235	<.001

Table 1: Summary of the fixed effects for a binomial regression model with lexical form use as the binary dependent variable, and by-participant and by-item random effects. Bold: The coefficients corresponding to the predicted effects (i)-(ii).

presence of a formally similar distractor sentence.

	+ distractor	- distractor
compositional form	14.76%	0.26%
lexical form	0.62%	0.18%

Table 2: Proportion of ‘force errors’ in Receiver trials

Moreover, our data confirm that (shape) ambiguity is more likely to be disambiguated in the [+part] conditions. To investigate whether disambiguation is a predictor of successful communication, we defined a measure of (strength of) convergence as shown in (4). For each pair of participants, convergence is defined as the deviation from chance in the probability of using the lexical nex form for a frequent $\neg\exists$ picture. Convergence ranges from 0, if a pair chooses lexical and compositional forms inconsistently, to 1, if a pair chooses only lexical or only compositional nex for the frequent objects.

$$(4) \text{ Convergence} = |p - 0.5| * 2$$

where $p = P(\text{lexical nex} \mid \text{+frequent } \neg\exists \text{ picture})$

As visualized in Fig. 4, convergence is a reliable predictor of the final score achieved by the pairs in the [+part] conditions. A linear model fitted to the final scores with convergence as predictor reveals a significant effect ($\beta = 67.1$, $t = 7.0$, $p < .001$, 95% CI [47.7, 86.5]): pairs with stronger convergence achieve higher scores. While this is true regardless of whether pairs converge on the compositional or the lexical form for frequent shapes, Fig. 4 also suggests that convergence on lexical nex is more common than convergence on compositional nex, and associated with higher strength of convergence. To test this, we use *signed convergence* (calculated as $p * 2 - 1$) as a measure of strength of convergence that accounts for convention direction. A one-sample t-test shows that mean signed convergence across pairs differs significantly from zero (mean = 0.4, $t = 4.46$, $p < .001$, 95% CI [0.22, 0.59]). Positive mean signed convergence indicates an overall bias toward convergence on the lexical nex form for frequent $\neg\exists$ pictures, consistent with our overall hypothesis.

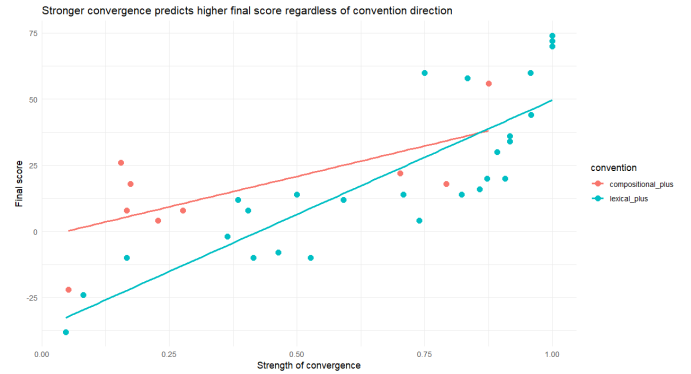


Figure 4: Score of each pair against strength of convergence in [+part] conditions. Blue: pairs that used more lexical than compositional forms in frequent $\neg\exists$ trials. Red: pairs that used more compositional forms in frequent $\neg\exists$ trials.

Conclusions

Our dyadic communication experiment provides direct experimental evidence that communicative accuracy and cognitive effort jointly shape lexicalization, biasing speakers toward using lexical forms for more frequent meanings over compositional alternatives. By extending the logic of the Principle of Least Effort beyond word length, our results show that efficiency considerations operate at the level of mappings between syntactic forms and meanings. In doing so, the study offers a novel experimental bridge between classic efficiency-based accounts of Zipf’s Law of Abbreviation and broader cross-linguistic patterns in lexicalization (e.g. Khishigsuren et al., 2025; Xu et al., 2024), supporting PLE as a general organizing principle of semantic systems.

Our proposal that cognitive effort shapes lexicalization is fully consistent with previous work suggesting that compositionality emerges from joint pressures for communicative accuracy and compressibility (e.g., Kirby et al., 2015). These views are complementary, addressing distinct levels of linguistic organization. While Kirby et al. (2015) contrast purely holistic (entirely lexical) languages with compositional ones, our study evaluates a finer-grained distinction within compositional systems—specifically, the optimal distribution of forms based on frequency (i.e., mapping frequent meanings to lexical forms and infrequent meanings to compositional forms). Furthermore, the artificial languages in Kirby et al.’s paradigm lack a counterpart to our distractor form, which makes order-sensitivity and structural derivation highly salient. It is thus likely that the processing-based cognitive effort associated with compositional forms does not manifest as significantly in their setup. Ultimately, learnability and compressibility pressures may drive the overarching emergence of a compositional system, while processing-based cognitive effort fine-tunes it, dictating the strategic distribution of lexicalized forms for highly frequent meanings.

Acknowledgments

We are grateful to our RA James Engels, who coded earlier versions of the experiment. We thank Jennifer Culbertson, Kenny Smith, Simon Kirby and audiences at the Mayfest (University of Maryland, May 2025), Waseda University Language Science Colloquium (February 2026), and UNPAG Kick-off Workshop (University of Göttingen, March 2026) for helpful discussions. This research was supported by UKRI Future Leaders Fellowship (MR/V023438/1; APP75291).

References

- Aylett, M., & Turk, A. (2004). The smooth signal redundancy hypothesis: A functional explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech. *Language and speech*, 47(1), 31–56.
- Bell, K., Brandt, S., Lieven, E., & Theakston, A. (2024). The acquisition of english modal constructions: A corpus-based analysis. *Journal of child language*, 51(5), 1134–1171.
- Denić, M., Steinert-Threlkeld, S., & Szymanik, J. (2022). Indefinite pronouns optimize the simplicity/informativeness trade-off. *Cognitive Science*, 46(5), e13142.
- Ferrer-I-Cancho, R., & Solé, R. V. (2003). Least effort and the origins of scaling in human language. *Proceedings of the National Academy of Sciences*, 100(3), 788–791.
- Imel, N., Guo, Q., & Steinert-Threlkeld, S. (2026). An efficient communication analysis of modal typology. *Open Mind*, 10, 1–28.
- 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.
- Kemp, C., Xu, Y., & Regier, T. (2018). Semantic typology and efficient communication. *Annual Review of Linguistics*, 4(1), 109–128.
- Khishigsuren, T., Mollica, F., Vylomova, E., & Kemp, C. (2025). Usage frequency predicts lexicalization across languages. *Proceedings of the 47th Annual Conference of the Cognitive Science Society*.
- Kirby, S., Tamariz, M., Cornish, H., & Smith, K. (2015). Compression and communication in the cultural evolution of linguistic structure. *Cognition*, 141, 87–102.
- Levshina, N. (2022). *Communicative efficiency*. Cambridge University Press.
- Mucha, A., Engels, J. J., Whibley, F., & Uegaki, W. (2024). Negative modality in Hausa, Thai and Kĩtharaka. *Proceedings of TripleA 10*, 1–17.
- Simpson, A. (2001). Focus, presupposition and light predicate raising in East and Southeast Asia. *Journal of East Asian Linguistics*, 10(2), 89–128.
- Uegaki, W., Mucha, A., Hannon, E., Engels, J., & Whibley, F. (2025). Cross-linguistic dataset of force-flavor combinations in modal elements. *Linguistic Variation*, 25(2), 269–305.
- Van Dooren, A., Dieuleveut, A., Cournane, A., & Hacquard, V. (2022). Figuring out root and epistemic uses of modals: The role of the input. *Journal of Semantics*, 39(4), 581–616.
- Veselinovic, D. (2019). *The syntax and acquisition of modal verb flavors* [Doctoral dissertation, New York University].
- Visonyanggoon, S. (2000). *Parallelism between noun phrases and clauses in Thai* [Doctoral dissertation, Michigan State University].
- Xu, A., Kemp, C., Frermann, L., & Xu, Y. (2024). Word reuse and combination support efficient communication of emerging concepts. *Proceedings of the National Academy of Sciences*, 121(46), e2406971121.
- Zipf, G. K. (1936). *The psychobiology of language*. Routledge.
- Zipf, G. K. (1949). *Human behavior and the principle of least effort: An introduction to human ecology*. Addison-Wesley.