Syntactic Choice Is Shaped by Fine-Grained, Item-Specific Knowledge

Emily Goodwin (goodwine@stanford.edu)
Department of Linguistics
Building 460 Stanford, CA 94305 USA
Beth Levin (bclevin@stanford.edu)
Department of Linguistics
Building 460 Stanford, CA 94305 USA
Emily Morgan (eimorgan@ucdavis.edu)
Department of Linguistics
469 Kerr Hall Davis, CA 95616 USA

Abstract

There is a longstanding debate over how much idiosyncratic, item-specific knowledge is contained in our mental grammars, in addition to productive knowledge of item-general rules and constraints. A key source of evidence is ordering preferences for syntactic alternations like the dative ("throw him an apple" vs. "throw an apple to him") vary depending on which words they contain. But the quantitative extent of this variability is poorly understood, especially in relation to superficially similar, non-dative constructions which are not alternating ("throw the man to the floor" vs. "*throw the floor the man"). To address this, we built a large corpus of naturally-occurring sentences including either dative or superficially similar non-dative structures, and analyzed the unique contributions of productive and verb-specific knowledge in predicting argument ordering preferences.

Keywords: syntactic alternations; exemplar models

Introduction

To what degree do speakers rely on productive knowledge of their language, and to what degree do they draw on their direct experience with specific words? Consider the ordering alternatives for binomial phrases, like *salt and pepper* vs. *pepper and salt*. Speakers have detailed knowledge of the preferred ordering of such expressions; in this case, their item-specific experience consists of previous exposures to either ordering, and their productive knowledge consists of generally applicable rules of semantics and phonology (for example, that longer constituents are preferred later).

Researchers agree that speakers use some amount of item-specific knowledge, but disagree about how much and under what conditions it is recruited. Theories vary widely: the words-and-rules theory predicts that item-specific knowledge is reserved for rare cases such as irregular forms, idioms, and extremely high frequency items (Pinker & Ullman, 2002), while usage-based theories claim that speakers store their lifetime of linguistic experience as imperfect memory traces, or *exemplars* (Ambridge, 2020; Bybee & Eddington, 2006; Bybee, 2006; Bybee & McClelland, 2005; Jackendoff, 2002; Goldberg, 2016; Jackendoff, 2007).

Wide-coverage quantitative evaluations of these claims have been lacking until recently (but see O'Donnell et al., 2009). Work on ordering preferences in binomials has shown that direct experience and productive knowledge trade-off to predict speaker ordering preferences. More frequent items rely more on direct experience; for most binomials, speakers use both their productive knowledge and their direct experience to varying degrees (Morgan & Levy, 2016, 2024).

	DO Form NP NP Structure	*to*- Form NP *to* NP Structure
"Dative Use" Has recipient	*Take him a cup*	*Take a cup to him*
"Non Dative Use" No recipient	*Take me a day*	*Take it to the limit*

Table 1: Dative and non-dative uses of the dative verb *take*.

However, it remains unclear whether the frequency-mediated trade-off between direct experience and productive knowledge is an idiosyncratic property of binomials. The question is complicated by linguistic knowledge of abstractions like syntactic constructions, which lack clear criteria for what learners consider relevant experience. This is especially crucial for models of statistical learning which rely on indirect negative evidence, leading to the entrenchment vs. preemption debate (Ambridge et al., 2014, 2018; Brooks & Zizak, 2002; Goldberg, 2011).

To see the problem, consider the dative alternation, exemplified in the first row of Table 1. Dative verbs convey the transfer of some object (the *theme* argument) to someone or something (the *recipient*), and permit different orderings of the arguments: the recipient appears first in the *Double-Object* (DO) form, and the theme appears first in the *to*-form. For a given verb and pair of arguments, speakers have preferences over which form is preferred. Bresnan et al. (2007) showed that these ordering preferences are in part captured by productive constraints on the arguments' phonology (for example, shorter arguments are preferred earlier) and semantics (for example, animate and definite arguments are preferred earlier). At the same time, Bresnan et al. also found that individual verbs still occur more or less frequently in a structure than would be predicted by the productive constraints alone. These verb-specific preferences are a different kind of item-specific information, since they apply to all uses of the verb, regardless of the chosen arguments. For example, both *give him the letter* and *give the woman an apple* count as direct experience of the verb *give*; both influence speakers' preferences when ordering a new sentence, like *give me a break* (Bridgwater et al., 2019). This contrasts with exist-

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

ing work on item-specific experience, where speaker's direct experience is conceptualized straightforwardly as every instance they have heard of the entire phrase (*salt and pepper* or *pepper and salt*). This difference motivates our first question: do abstract grammatical alternations like the dative display a trade-off between productive and item-specific knowledge?

A second question arises when considering item-specific experience at higher levels of abstraction: multiple constructions often have the same syntax (their constituents are configured in the same way) but different semantic representations (the constituents play different roles in the event). For example, all uses of *take* in Table 1 have two arguments; however, only those with a recipient ("dative uses") alternate (**take a day to me, *take the limit it*) (Rappaport Hovav & Levin, 2008). As we show below, the relative proportion of DO and *to*-forms differs in the dative and the non-dative uses of many verbs. For exemplar models and other usage-based theories, exemplars are explicitly form-meaning pairs (Goldberg, 2003); this could predict that speaker ordering preferences for dative uses are not influenced by non-dative uses of the same verb, even if the non-dative use is more frequent. This is our second question: do syntactically similar structures influence ordering preferences of the dative, and if so, how? We investigate the relative frequencies of dative to non-dative uses, so that non-dative structures are taken into account in the larger empirical picture.

Finally, frequency may play a different role in the trade-off between item-specific and productive knowledge when speakers have dramatically more or less experience with the items. By investigating verb-specific effects, we test a wider range of the item-frequency spectrum than other structures: while the most frequent binomial in Morgan & Levy (2015)'s corpus is estimated to be heard 7,300 times by the time a speaker reaches college-age (once per day), our most common verb (*give*) occurs about 163,371 times in the same time frame (22 times a day). Our least-frequent verb (*telephone*) is similar in frequency to the least-common binomial.

In this paper, we introduce and analyze a corpus of dative and non-dative examples[1]. We find that speakers recruit direct experience with verbs to structure their utterances, abstracting over individual examples with different arguments. As with binomials, more frequent verbs rely more on verb-specific knowledge. We also find that verb-specific preferences are only predicted by experience with true dative uses. Finally, we test our model on a dataset of human preference judgments, and find the same patterns in this dataset. These results suggest that speakers draw on experience with a particular verb across different uses, but distinguish between their experience of dative and non-dative uses.

Methods

To create the corpus, we first parsed a corpus of webtext, and extracted instances of dative verbs with two arguments. We then iteratively sampled instances of each verb and hand-annotated them, both for status as a dative or non-dative (collapsing together spatial goals, resultatives, and other non-dative examples) and for the productive constraints known to influence word ordering (described in more detail below).

Because frequency influences speakers' use of item-specific information in other structures, our corpus creation process emphasized representing the full spectrum of dative-verb frequency. We parsed 6.15 billion words of the English portion of the colossal cleaned Common Crawl (Raffel et al., 2020), equivalent to 2.39 times the lifetime language exposure estimated in Levy et al. (2012). This version of Common Crawl predates chatGPT and the machine-generated text currently on the web.

To identify datives, we first parsed the text using automatic dependency parser Stanza (Qi et al., 2020). We then extracted all sentences with a verb which was (1) listed as dative alternating in Levin (1993), and (2) took two arguments in either of the dative structures. We excluded examples with intervening material between the two arguments, like *throw [a pass] right/directly/straight [to the quarterback]*, as they influence ordering preferences for independent reasons (**throw right the quarterback a pass*).

We sampled and annotated 50 examples of each verb lemma, but found wide variation among verbs in the ratio of dative to non-dative uses. To ensure the dataset included enough dative uses of each verb, we annotated additional examples sampled non-uniformly (sampling more examples of verbs which showed dative use less frequently). Each example was annotated by two trained undergraduate annotators. If both annotators agreed the sentence was dative, the arguments were annotated for the properties relevant to productive rules. If the two annotators disagreed on any of the argument annotations, we recruited a third annotator; if the third judgment did not break the tie (for properties with more than two possible labels), the example was discarded.

We annotated the arguments for the following properties, used by Bresnan et al. (2007):

Givenness. Themes and recipients which referred to something already mentioned in the context were marked as given; otherwise, they were marked non-given. The context was defined as anywhere earlier in the target sentence, or in the two preceding sentences.

Animacy. Themes and recipients which referred to a human or animal were marked as animate, others (including organizations) were treated as inanimate.

Concreteness. Themes which referred to a thing which is prototypically concrete, meaning it can be interacted with using the five senses, were marked concrete. This property was added to compensate for the sparsity of theme animacy (see next section).

Pronominality. Whether the argument was a pronoun (annotated via automatic string-matching).

Definiteness. Whether the argument was grammatically definite or indefinite.

[1] Corpus available at https://github.com/emilygoodwin/LCOD.git
Code at https://github.com/emilygoodwin/dativeCorpusAnalyses.git

Number. Our parsing pipeline automatically assigned number as Singular, Plural, or NA. 13.2% of recipients and 1.5% of themes were marked NA, mostly driven by the pronoun *you* which is ambiguous for number in English.

Person. Whether the theme and recipient were local (first or second person) or non-local (third person).

Length Difference. A sign-preserving log transform of the difference in length between the theme and recipient, measured in graphemic words, plus a 0.5 smoothing coefficient to avoid taking the log of 0.

Verb sense. The meaning of the verb in context, according to the verb senses defined in Bresnan et al. (2007): Abstract, Transfer of Possession, Future Possession, and Communication. A fifth class, Prevention of Possession, (for verbs like *cost, deny*) does not appear in our dataset because none of the verbs listed as alternating by Levin (1993) have this sense.

Structural Parallelism. This variable captures the presence of other DO and *to*-forms in the two sentences preceding the dative example. Defined as 0 if there were no examples of DO or *to*-form, 1 if the last one in the context was DO, and -1 otherwise.

Results

The final corpus had $23,320$ sentences, including $7,278$ dative uses. Of the datives, 372 could not be tiebroken on some field; we included these in estimating the proportion of dative uses, but excluded them when fitting later models. We excluded 20 additional examples with complex themes or recipients where the definiteness of the entire constituent unclear (*the cat and a ball*). In the remaining analyses, we also limited our dataset to verbs which had 10 or more dative instances. This yielded a total of $6,837$ examples for 91 verbs. We met or exceeded our original goal of collecting 50 true dative examples for 68 verbs.

Overall, the *to*-form dominated both dative and non-dative examples in the corpus (Figure 1A). However, the degree of preference for the DO form varied among verbs: measured over true datives, the proportion of DO instances ranged from .98 (*tell*) to none at all (18 verbs, typically conveying motion like *cart* and *drive*). We plot verbs' preferences for the DO form in Figure 1C, in the x-axis of the right panel.

For most verbs, the proportion of instances in the DO form was higher among dative than non-dative uses (Figure 1C); together with the relative frequency of non-datives, this suggests that speakers are exposed to plenty of *to*-forms in non-dative uses. Whether they draw on this evidence is discussed in the next section.

We also found that preference for the DO form was predicted by verb frequency (Figure 1B). A model predicting each verb's preference for the DO form from its log-transformed frequency found a significant effect of frequency ($\beta = 0.04$, $S.E. = 0.01$, $P < .001$). One possible explanation is that the DO form, being less frequent than the *to*-form, requires more exposure to learn. As a consequence, only highly frequent verbs yield enough evidence to learn this structure.

Modeling Productive and Verb-Specific Knowledge

The variation in DO preference among verbs may be a result of applying the same productive constraints to verbs with different semantics: because different verbs have different meanings, they may appear more or less frequently with particular themes and recipients, which would drive them to appear in different forms as a function of the productive constraints. For example, *give* may occur frequently in the DO form because it occurs frequently with pronominal recipients. In an extreme case, there would be no need to model speakers' direct experience with individual verbs: the productive constraints alone might predict highly skewed preferences.

To assess the independent contribution of verb-specific experience, we fit a Bayesian logistic model using BRMS (Bürkner, 2017) on the dative uses in our corpus. The dependent variable was the form (DO or *to*-form). The independent variables were the annotation properties introduced above. Two theme predictors were sparse: animacy (only 4.14% of themes are animate) and grammatical person (0.32% of themes are not in third person); following Bresnan et al. (2007), we removed these predictors and added theme concrete to compensate. All the other argument predictors (givenness, pronominality, definiteness, and number) were included for both theme and recipient. Recall that we were interested in the amount of information that can be attributed to direct experience with a verb: for this reason, we included a by-verb random intercept term. Unlike Bresnan et al. (2007), this was not crossed with verb sense: while Bresnan et al.'s model fit one intercept for each verb in each sense, we fit just one intercept for all uses of the verb. The estimates for fixed effects associated with verb sense, plural recipients, plural and singular themes, and previous structures (PP or DO) had credible intervals which fell within a range including 0 (see Figure 2).

To distinguish between idiosyncratic verb preferences and preferences driven by the fixed effects, we extracted the model predictions for all the dative sentences in our corpus, based only on the fixed effects (including the overall intercept), excluding the random intercept. Figure 3A shows the predicted proportion of DO-forms among datives against the true observed preferences for each verb. The productive constraints failed to predict the distribution of observed preferences, missing many verbs with an extreme observed preference for the DO or *to*-form; this suggests that verb-specific knowledge helps drive those preferences.

Next, we asked whether frequency can predict idiosyncrasy. In Figure 3B, we plot the magnitude of the difference between the verbs' predicted ordering preference and its observed ordering preference, as a function of verb frequency. More frequent verbs were more likely to have ordering preferences that diverged more from those predicted by the productive constraints. We speculate this is because only more frequent verbs provide speakers with enough evidence to learn an ordering preference that departs from the productive constraints.

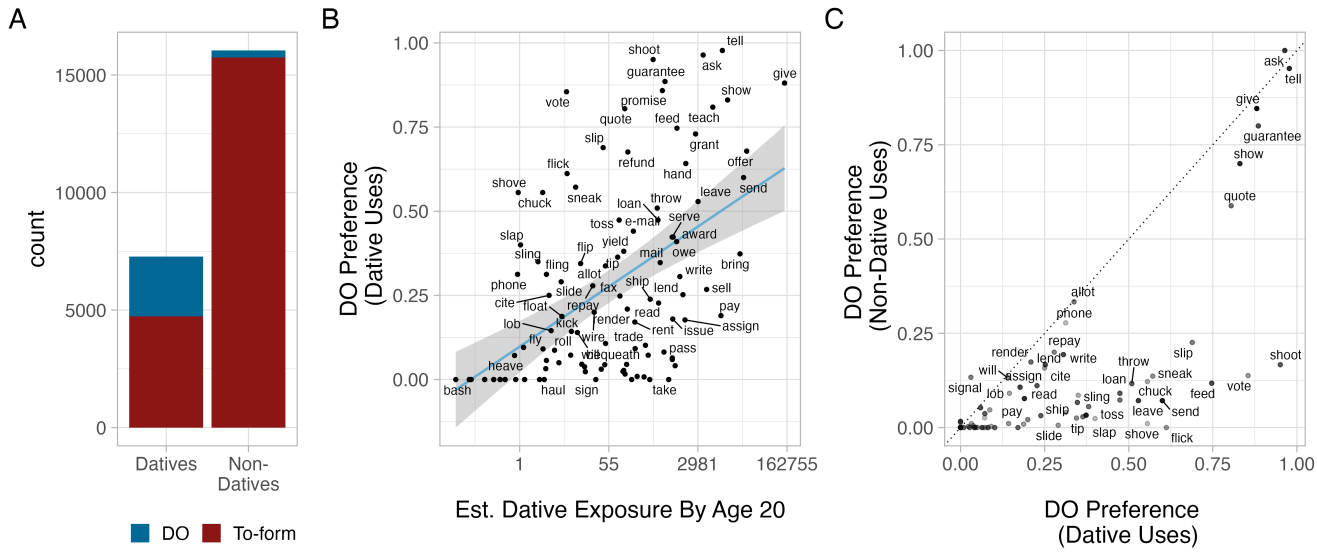

Figure 1: **A**: Proportion of structures in the corpus. **B**: The proportion of dative uses in the DO form for each verb, as a function of the estimated number of exposures to the verb in a dative use, by age 20, with the x-axis log scaled. This estimate relies on total language exposure estimates from Levy et al. (2012). **C**: The proportion of occurrences in the DO form as measured over dative uses (x axis) or over non-dative uses extracted for the corpus (y axis). Darkness of each point scales with verb frequency.

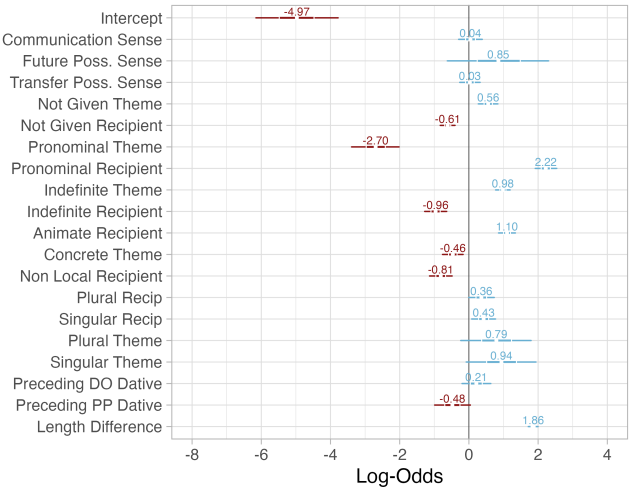

Figure 2: Untransformed estimated effect sizes. The model's dependent variable (syntactic form) had reference level *to*-form, contrast level DO form.

Superficially Similar Structures Do Not Influence Dative Ordering

Recall that verb preferences for the DO structure differ when considering dative uses only, non-dative uses only, or both; one of our goals was to identify whether speakers draw on their experience of non-dative uses when choosing how to order a dative.

To determine whether non-dative uses influence speaker preferences in the dative alternation, we extracted the by-verb random intercepts from our earlier model (which describe verb-specific preferences for either form, independent of productive constraints). We then fit a linear regression, predicting the by-verb effect from two fixed effects: the proportion of examples in the DO structure for each verb, as measured over only dative uses; and the proportion as measured over non-dative uses. An increased preference for the DO use among dative uses was predictive of the verb-specific preferences ($\beta = 6.01$, $S.E. = 0.47$, $P < .001$); increased preference for the DO structure among non-dative was not ($\beta = 0.63$, $S.E. = 0.53$, $P = 0.242$). This suggests that speakers draw on their direct experience with dative uses of a verb; their experience with non-dative uses did not have a detectable effect.

Comparison With Human-Reported Preferences

We also tested our model on a dataset of human preference judgments for 2,050 dative sentences collected by Hawkins et al. (2020). In this dataset, participants indicated on a slider how strongly they preferred the *to*-form (0) or the DO form (1), with the midpoint labeled as "about the same". Each verb was shown with five different recipient arguments (which varied in length and definiteness, but were all animate) and five theme arguments (which varied in definiteness but were all three words or shorter).

The human data represent a different assessment of ordering preferences, both as a consequence of how they are collected (introspective judgments rather than corpus frequencies), and as a consequence of the arguments considered (participants judged only animate recipients and a small set of themes). This makes it difficult to directly test our model, which was fit to corpus frequencies, against the human pref-

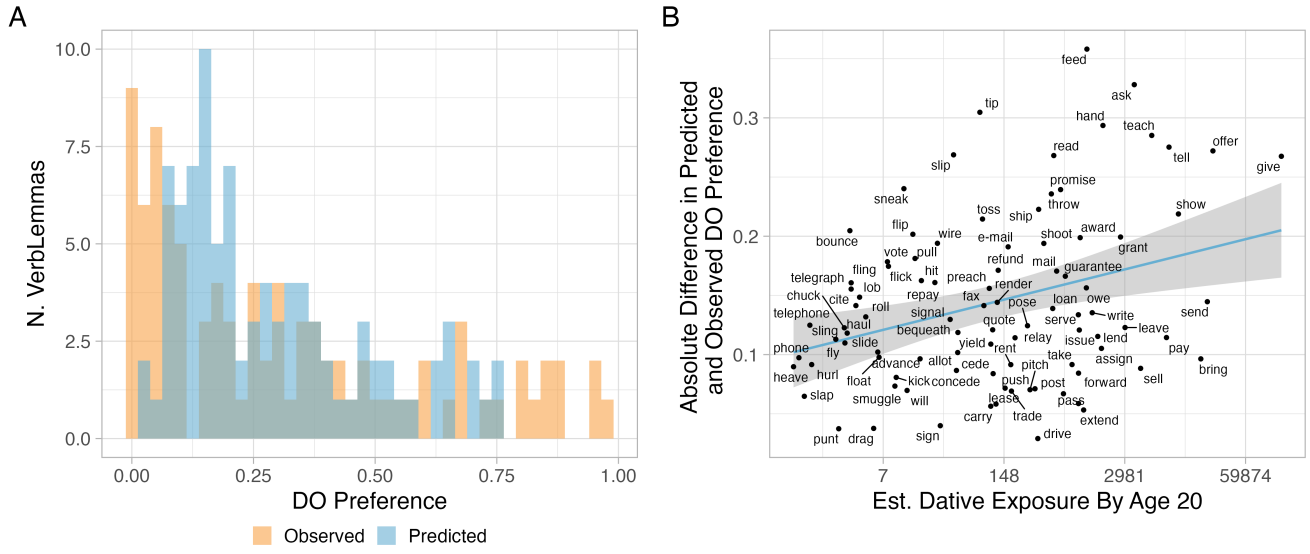

Figure 3: **A:** The difference between observed DO preference and predicted DO preference. Predictions are generated without the model's random effect structure, reflecting only productive knowledge. The darkest shaded areas are overlaps in the two distributions. **B:** Verb idiosyncrasy (mean absolute difference between the predicted and observed DO preference) as a function of verb frequency (estimated exposure to the verb in the dative structure, by age 20, with the x-axis log-scaled). Model predictions were generated without random effect structure, and represent speakers' productive knowledge.

erences dataset. Nevertheless, we can test whether the qualitative patterns we observed in the corpus also hold for the human dataset: First, the pattern that more frequent verbs will be more idiosyncratic and preferred in the DO form; second, that the distribution of structures among dative uses (but not non-dative uses) will be predictive of ordering preferences.

We extracted predictions for the sentences in the human-preferences dataset from our model. Recall that the model predicts syntactic form (DO or *to*-form) from fixed-effects reflecting productive knowledge and random by-verb intercepts reflecting direct experience. Figure 4A shows the distribution of predicted and observed human ordering preferences. Human-reported preferences were approximately normally distributed, with a mean closer to .5 than the model; the model predictions were, as with the corpus data, highly skewed toward the *to*-form. The Pearson correlation between the predicted DO-preference for each verb, and its mean reported DO-preference, was 0.76.

To test whether more frequent verbs are more idiosyncractic, we plot the difference between the mean model prediction based only on fixed effects and the mean human preference for each verb in Figure 4B. As with the corpus data, more frequent verbs had more idiosyncratic ordering preferences. We also found that more frequent verbs were associated with higher reported preferences for the DO form ($\beta = 0.02$, $S.E. = 0.003$, $P < .001$).

Finally, we tested whether the human preferences were better predicted by the distribution of forms among dative uses of a verb, or both dative and non-dative uses. We fit a model predicting the mean human ordering preference for each verb from the observed ordering preference in the corpus of all datives, non-datives, and the model predictions (calculated as the mean over the predictions from fixed effects only). Only the preference over dative uses was predictive ($\beta = 0.20$, $S.E. = 0.04$, $P < .001$), and there was no effect of non-dative uses ($\beta = 0.04$, $S.E = 0.05$, $P = .42$) or model predictions ($\beta = .71$, $S.E = 0.44$, $P = 0.11$).

Although the human dataset assesses ordering preferences in a different way than corpus frequencies, we found the same qualitative patterns: strong by-verb variability, with more idiosyncratic verbs being more frequent; more frequent verbs favoring the DO form; and the form preference measured over just datives being more predictive than over other structures.

Discussion

Verb Idiosyncracy and Frequency. While productive knowledge identified by prior work captures much variation in the argument-ordering preferences of dative structures, direct experience with verbs also plays a crucial role, with more frequent verbs having more idiosyncratic ordering preferences. This pattern suggests that, in dative productions, the extent to which speakers draw on productive vs. verb-specific knowledge is mediated by the verb's frequency.

Frequency and The DO Form. We also found that increased verb frequency was associated with a preference for the DO form. This could be explained as a pressure to regularize: since the *to*-form is more common (among both dative and non-dative uses), speakers may treat it as a default while the less-common DO structure is available for only the most fre-

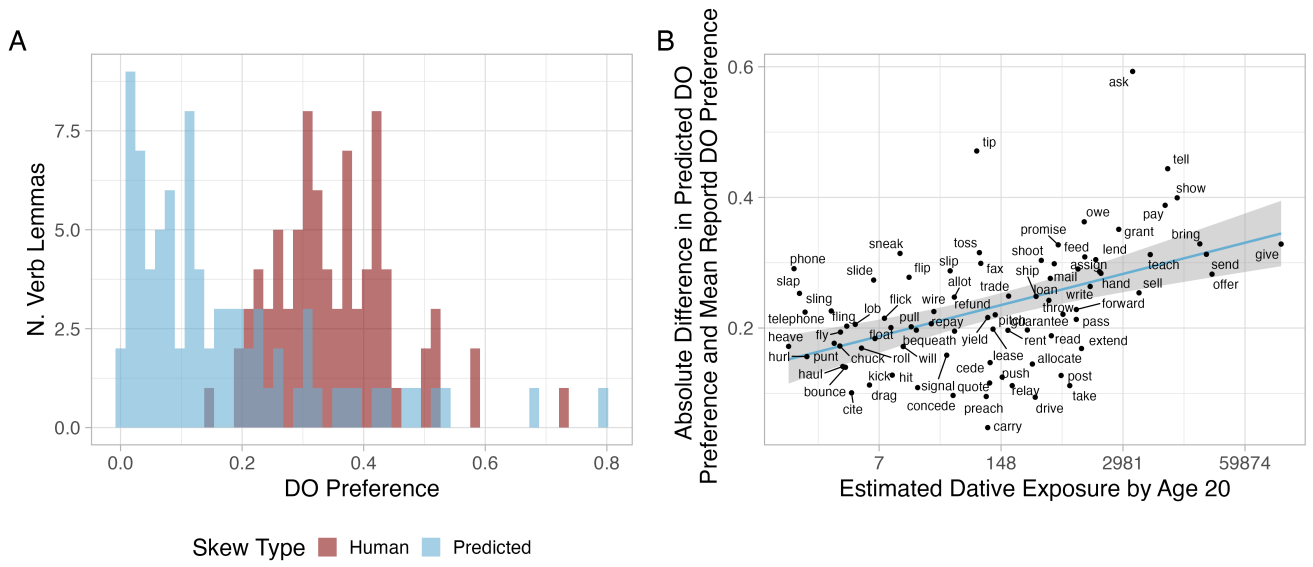

Figure 4: **A**: Verbs' human-reported preference judgements against the predictions of our corpus model. Model predictions were generated including the random effect structure, and represent speakers' productive and item-specific knowledge. **B**: Verb idiosyncrasy (mean absolute difference between proportion of predictions in DO form and mean slider rating) against frequency (estimated exposure to the verb in the dative structure, by age 20, with intervals natural-log transformed). Model predictions were generated without random effect structure, and represent speakers' productive knowledge.

quent verbs. Together with the observation that more frequent verbs are more idiosyncratic, this proposal would be similar to comparable proposals in work on morphological irregularity (Wu et al., 2019). However, note that word frequency correlates with a number of other semantic factors (Piantadosi et al., 2012; Winter & Srinivasan, 2022); we do not rule out the possibility that these could cause the increase in DO forms.

The Role of Non-Dative Uses. Finally, we investigated whether speakers' ordering preferences for a verb are influenced by their experience with the verb in superficially similar, non-dative uses. We found that verbs' distribution of DO and *to*-form structures in non-dative uses did not influence ordering preferences. This suggests that the fact that lower-frequency verbs are more likely to prefer the *to*-form is not simply a function of how often they appear in non-dative *to*-forms. This is consistent with two explanations: Non-dative experience may be completely irrelevant for speakers when structuring dative uses. Or, non-dative experience could contribute to learning the *to*-form as globally acceptable, without contributing to a verb-specific preference.

Why would speakers not consider all the experience they have with a verb? Under a theory where constructions are represented as form-meaning pairs, the meaning part of the construction could include a characterization of the event participants (distinguishing between recipients and spatial goals, or some other semantic variable which distinguishes between datives and non-dative senses). This is in line with some previous work on syntactic exemplar models (Bybee & Eddington, 2006; Bybee, 2006). To predict our findings, the analogy mechanism used by speakers during production would need to be sensitive to the coding of event participants in the exemplar, weighting non-dative examples less highly.

Conclusion

In this paper, we investigated the contributions of productive knowledge and direct experience to speakers' productions of the dative alternation. We built a large corpus of naturally-occurring sentences including either dative or superficially similar non-dative structures, and analyzed the unique contributions of productive and verb-specific knowledge in predicting argument ordering preferences.

We found that individual verbs have unique ordering preferences not predicted by the productive constraints, with more frequent verbs having stronger idiosyncratic preferences and preferring the DO form overall. This suggests that speakers abstract over their exposure to the same verb with different complements, relying more on this knowledge for frequent verbs. At the same time, speakers do not abstract over all their experience with a verb: ordering preferences for datives are not influenced by non-dative constructions, even when the non-dative constructions use the same verb and surface syntax, and are more frequent than dative uses.

Finally, we applied our model to a corpus of human-reported ordering preferences, and found the same qualitative patterns. Our results support both (1) a substantial role for item-specific knowledge in shaping human syntactic choice and (2) fine-grained distinctions between items in the mental grammar, such that experience with one item may have little influence on others with superficially similar form.

Acknowledgments

This project was funded by NSF#2334255 and UC Davis College of Letters and Science. We are grateful to Robert Hawkins, Cory Shain, the SoIL and CLiMB labs; to the anonymous CogSci reviewers for their very helpful feedback; and to the undergraduate annotators who contributed to this work: Ian Miranda, Luna Llamas, Heidi Trinh, Joseph Bender, Ulisses Gallardo, Ariel Padovitz, Isabella Xu, Ellie Bi, Jay Simpson, Aquarius Wong, Alejandra Mercado, Meghana Kotha, Maya Hill, Laasya Babbellapati, and Edria Jabil.

References

Ambridge, B. (2020). Against stored abstractions: A radical exemplar model of language acquisition. *First Language*, *40*(5-6), 509–559.

Ambridge, B., Barak, L., Wonnacott, E., Bannard, C., & Sala, G. (2018). Effects of both preemption and entrenchment in the retreat from verb overgeneralization errors: Four reanalyses, an extended replication, and a meta-analytic synthesis. *Collabra: Psychology*, *4*(1).

Ambridge, B., Pine, J. M., Rowland, C. F., Freudenthal, D., & Chang, F. (2014). Avoiding dative overgeneralisation errors: Semantics, statistics or both? *Language, Cognition and Neuroscience*, *29*(2), 218–243.

Bresnan, J., Cueni, A., Nikitina, T., & Baayen, R. H. (2007). Predicting the dative alternation. In *Cognitive foundations of interpretation* (pp. 69–94). KNAW.

Bridgwater, E., Kyröläinen, A., & Kuperman, V. (2019). The influence of syntactic expectations on reading comprehension is malleable and strategic: An eye-tracking study of English dative alternation. *Canadian Journal of Experimental Psychology*, *73*(3), 179–192.

Brooks, P. J., & Zizak, O. (2002). Does preemption help children learn verb transitivity? *Journal of Child Language*, *29*(4), 759–781.

Bybee, J. (2006). From usage to grammar: The mind's response to repetition. *Language*, *82*(4), 711–733.

Bybee, J., & Eddington, D. (2006). A usage-based approach to Spanish verbs of 'becoming'. *Language*, *82*(2), 323–355.

Bybee, J., & McClelland, J. L. (2005). Alternatives to the combinatorial paradigm of linguistic theory based on domain general principles of human cognition. *The Linguistic Review*, *22*(2-4).

Bürkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, *80*(1), 1–28.

Goldberg, A. E. (2003). Constructions: A new theoretical approach to language. *Trends in Cognitive Sciences*, *7*(5), 219–224.

Goldberg, A. E. (2011). Corpus evidence of the viability of statistical preemption. *Cognitive Linguistics*, *22*(1), 131–153.

Goldberg, A. E. (2016). Partial productivity of linguistic constructions: Dynamic categorization and statistical preemption. *Language and Cognition*, *8*(3), 369–390.

Hawkins, R., Yamakoshi, T., Griffiths, T., & Goldberg, A. (2020). Investigating representations of verb bias in neural language models. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (pp. 4653–4663). Association for Computational Linguistics.

Jackendoff, R. (2002). What's in the lexicon? In S. Nooteboom, F. Weerman, & F. Wijnen (Eds.), *Storage and computation in the language faculty.* Dordrecht: Springer Netherlands.

Jackendoff, R. (2007). A Parallel Architecture perspective on language processing. *Brain Research*, *1146*, 2–22.

Levin, B. (1993). *English verb classes and alternations: A preliminary investigation*. The University of Chicago Press.

Levy, R., Fedorenko, E., Breen, M., & Gibson, E. (2012). The processing of extraposed structures in English. *Cognition*, *122*(1), 12–36.

Morgan, E., & Levy, R. (2015). Modeling idiosyncratic preferences: How generative knowledge and expression frequency jointly determine language structure. In *Proceedings of the 37th annual conference of the cognitive science society.*

Morgan, E., & Levy, R. (2016). Abstract knowledge versus direct experience in processing of binomial expressions. *Cognition*, *157*, 384–402.

Morgan, E., & Levy, R. (2024). Productive knowledge and item-specific knowledge trade off as a function of frequency in multiword expression processing. *Language*, *100*(4), e195–e224.

O'Donnell, T. J., Tenenbaum, J. B., & Goodman, N. D. (2009). *Fragment grammars: Exploring computation and reuse in language*. MIT Computer Science and Artificial Intelligence Laboratory Technical Report Series.

Piantadosi, S. T., Tily, H., & Gibson, E. (2012). The communicative function of ambiguity in language. *Cognition*, *122*(3), 280–291.

Pinker, S., & Ullman, M. T. (2002). The past and future of the past tense. *Trends in Cognitive Sciences*, *6*(11), 456–463.

Qi, P., Zhang, Y., Zhang, Y., Bolton, J., & Manning, C. D. (2020). Stanza: A Python natural language processing toolkit for many human languages. In *Proceedings of the 58th annual meeting of the association for computational linguistics: System demonstrations* (p. 101-108).

Raffel, C., Shazeer, N., Roberts, A., Lee, K., Narang, S., Matena, M., ... Liu, P. J. (2020). Exploring the limits of transfer learning with a unified text-to-text transformer. *Journal of Machine Learning Research*, *21*(1), 1–67.

Rappaport Hovav, M., & Levin, B. (2008). The English dative alternation: The case for verb sensitivity. *Journal of Linguistics*, *44*(1), 129–167.

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.

Wu, S., Cotterell, R., & O'Donnell, T. (2019). Morphological irregularity correlates with frequency. In *Proceedings of the 57th annual meeting of the association for computational linguistics* (pp. 5117–5126).