Frequency-Dependent Regularization in Syntactic Constructions

Zoey Liu
Department of Computer Science
Boston College
ying.liu.5@bc.edu

Emily Morgan
Department of Linguistics
University of California, Davis
eimorgan@ucdavis.edu

Previous research has suggested language users tend to have consistent ordering preferences given a syntactic construction with different grammatical alternatives (e.g. within the same binomial type, *theory and data* is consistently more preferred than *data and theory*). Borrowing the term of a well-known phenomenon in statistical learning, the tendency to make language structures more consistent is known as *regularization*. From the perspective of language learning and production, regularization means when there is variation in the input, language users will preferentially minimize the amount of variation in the output by reproducing the most frequent structure among all possible alternatives. For example, if a speaker has encountered *theory and data* 64 times and *data and theory* 36 times (a 1.8:1 ratio), regularization means the speaker will produce the former even more frequently than it was heard and thereby increases its output probability.

However, the tendency to regularize contradicts the dominant view from rational language comprehension (Levy, 2008), that since language users are sensitive to the probabilistic distributions of different structures, they will *probability match* rather than regularize. This means with a given construction as the input, language users will maintain the amount of variation of different alternatives in the output. For instance, with the same binomial type above, in production languages users will approximately maintain the 1.8:1 ratio (*theory and data* vs. *data and theory*) for the two structures as well.

Most experiments that tried to tease apart regularization and probability matching have focused on the learning and production of morphemes (Saldana et al., 2017), words (e.g. combining nouns with determiners) (Hudson Kam and Newport, 2005; Ferdinand et al., 2019; Smith et al., 2017), or word orders (Culbertson and Smolensky, 2012; Hudson Kam, 2019) *only* with artificially constructed stimuli. A few studies examined natural online production in sign languages, yet focusing mainly on morphological variation (e.g. Nicaraguan Sign Language (Senghas, 1995); American Sign language (Ross and Newport, 1996; Singleton and Newport, 2004). By contrast, explorations of regularization in syntactic constructions using naturalistic data have been lacking. Thus in general, whether language users tend to perform probability matching or regularization when reproducing structural variants and under what contextual conditions remain understudied.

Two notable exceptions thus far have attempted to narrow this gap, both using corpus data. Morgan and Levy (2016) demonstrated that the extent of regularization in binomial expressions in English is affected by the frequency of a binomial type: the higher the frequency, the stronger and more extreme preference there is for one alternative over the other (e.g. *safe and sound* $>>>$ *sound and safe*; *facts and techniques* $>$ *techniques and facts*). In other words, regularization is *frequency-dependent*. Liu and Morgan (2020) further demonstrated this regularization bias in the dative construction in English (Bresnan et al., 2007), showing that the preference extremity within a dative type depends on its type frequency. Taking a data-driven approach, this study builds upon previous work and addresses two questions: (1) how wide-spread does frequency-dependent regularization exist in different syntactic construction types? (2) if this regularization bias were to hold regardless of the particular syntactic constructions, then what are the explanatory motivations behind it?

Overall type frequency We used as test cases adjective orders in phrases of the form Adj-Adj-N (AAN) and the dative constructions in English, the latter considered as a higher-level construction type than the former and the binomial types previously tested. An AAN type consists of two alternatives where the order of the two adjectives is

Proceedings of the Society for Computation in Linguistics (SCiL) 2021, pages 387-389.
Held on-line February 14-19, 2021

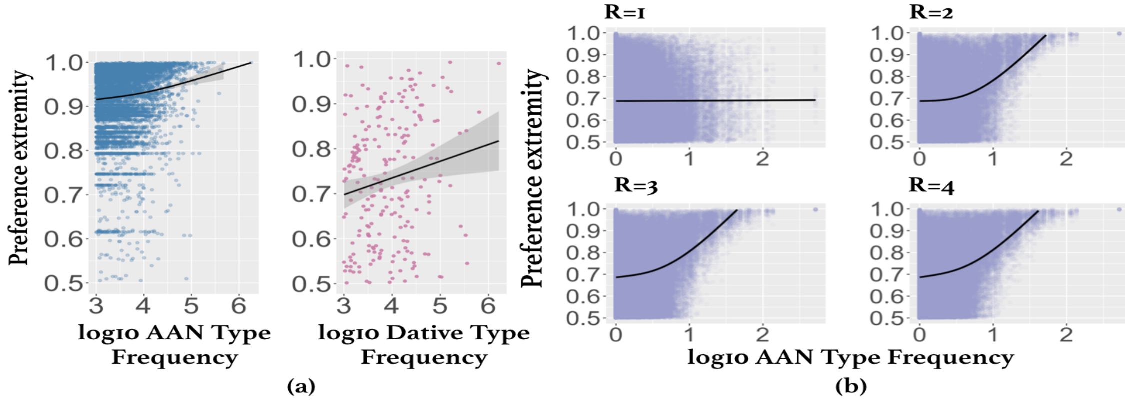

Figure 1: Selected results for AAN and datives with Levin class verbs: (a) demonstration of existence of frequency-dependent regularization (AAN: β = 0.02 (0.01, 0.02); dative: β = 0.04 (0.01, 0.07)); (b) simulating AAN corpus data using ILMs with different R (a free parameter in the frequency-independent regularization function). Frequency-dependent regularization emerges with R>1. All significance testing was from Bayesian linear regression predicting preference extremity as a function of overall type frequency.

switched yet with the same head noun (e.g. *little red corvette* vs. *red little corvette*). We took test data from Futrell et al. (2020) and selected AAN types with a type frequency N$\geq$1000 based on estimates from the Google n-gram corpus (n=8868). A dative construction can be realized as either the double object structure (V-NP-NP; *Black Panther **gave*** $[_{NP}$ *female characters]* $[_{NP}$ *the proper spotlight]*) or the prepositional object structure (V-NP-PP; *Black Panther **gave*** $[_{NP}$ *female characters]* $[_{PP}$ *the proper spotlight]*), and different dative types were distinguished based on the head verbs. We used raw data from the CoNLL 2017 Shared Task on multilingual parsing for extraction of dative constructions and estimation of type frequency, and extended Liu and Morgan (2020) with two other ways of verb (thereby dative types) selections and more careful heuristic filtering: (1) verbs with N$\geq$1000 that occur in at least one of the dative structures (n=563), following Liu and Morgan (2020); (2) verbs with N$\geq$1000 that also appear in both structures for at least 100 times (n=374); (3) verbs with N$\geq$1000 that also belong to the dative or the benefactive verb class as defined by Levin (1993) (n=223).

Preference extremity Preference extremity was approximated as the probability of the more preferred structure given a construction type. Note the matter of interest here is how each construction type (e.g. verb idiosyncrasy for the dative constructions) affects regularization. Since the ordering preference of a dative construction is governed by different abstract constraints besides the verb (Bresnan et al., 2007), the effects of these constraints need to be excluded to more precisely quantify the role of the verb. To do that, we used mixed-effect models to predict the order, including phrasal length, definiteness, pronominality as fixed effects, and the verb as a random effect (models were fit to each of the three subsets described above). The probability for each structure within a dative type was computed based on just the random effect of the verb. Preference extremity was measured similarly for AAN; the model included the information-theoretic constraints (Futrell et al., 2020) as fixed effects and each AAN type as a random effect.

Iterated learning To address the evolution of frequency-dependent regularization, we used iterated learning models (ILM), which computationally simulate the process of how language structures evolve through generations. The learning process of ILM is a process of Bayesian inference. We augmented standard ILM in the same way as Liu and Morgan (2020), where a frequency-independent regularization function was applied in the data generation stage.

Results Results (e.g. Figure 1) show that frequency-dependent regularization exists in both AAN and the dative constructions with all three selections of verbs, and appropriate combinations of parameter settings in ILM give rise to this regularization bias. This indicates that language users regularize to some extent in production, and that interactions between language production and the continuous process of cultural transmission could lead to frequency-dependent regularization.

References

Joan Bresnan, Anna Cueni, Tatiana Nikitina, and R Harald Baayen. 2007. Predicting the dative alternation. In *Cognitive foundations of interpretation*, pages 69–94. KNAW.

Jennifer Culbertson and Paul Smolensky. 2012. A bayesian model of biases in artificial language learning: The case of a word-order universal. *Cognitive Science*, 36(8):1468–1498.

Vanessa Ferdinand, Simon Kirby, and Kenny Smith. 2019. The cognitive roots of regularization in language. *Cognition*, 184:53–68.

Richard Futrell, William Dyer, and Greg Scontras. 2020. What determines the order of adjectives in English? Comparing efficiency-based theories using dependency treebanks. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 2003–2012, Online. Association for Computational Linguistics.

Carla L. Hudson Kam. 2019. Reconsidering retrieval effects on adult regularization of inconsistent variation in language. *Language Learning and Development*, 15(4):317–337.

Carla L. Hudson Kam and Elissa L. Newport. 2005. Regularizing Unpredictable Variation: The Roles of Adult and Child Learners in Language Formation and Change. *Language Learning and Development*, 1(2):151–195.

Beth Levin. 1993. *English verb classes and alternations: A preliminary investigation*. University of Chicago press.

Roger Levy. 2008. Expectation-based syntactic comprehension. *Cognition*, 106(3):1126–1177.

Zoey Liu and Emily Morgan. 2020. Frequency-dependent Regularization in Constituent Ordering Preferences. In *Proceedings of the 42nd Annual Meeting of the Cognitive Science Society*, page 2990–2996.

Emily Morgan and Roger Levy. 2016. Frequency-Dependent Regularization in Iterated Learning. In *The Evolution of Language: Proceedings of the 11th International Conference (EVOLANG11)*.

Danielle S Ross and Elissa L Newport. 1996. The development of language from non-native linguistic input. In *Proceedings of the 20th Annual Boston University Conference on Language Development*, volume 2, pages 634–645. Cascadilla Press Boston.

Carmen Saldana, Kenny Smith, Simon Kirby, and Jennifer Culbertson. 2017. Is the strength of regularisation behaviour uniform across linguistic levels? In *Proceedings of the 39th Annual Meeting of the Cognitive Science Society*, pages 1023–1028.

Ann Senghas. 1995. *Children's contribution to the birth of Nicaraguan Sign Language*. Ph.D. thesis, Massachusetts Institute of Technology.

Jenny L Singleton and Elissa L Newport. 2004. When learners surpass their models: The acquisition of american sign language from inconsistent input. *Cognitive psychology*, 49(4):370–407.

Kenny Smith, Amy Perfors, Olga Fehér, Anna Samara, Kate Swoboda, and Elizabeth Wonnacott. 2017. Language learning, language use and the evolution of linguistic variation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1711):20160051.