Priming vs. Inhibition of Optional Infinitival *to*

Robin Melnick
Pomona College
robin.melnick@pomona.edu

Thomas Wasow
Stanford University
wasow@stanford.edu

Abstract

The word *to* that precedes verbs in English infinitives is optional in at least two environments: in what Wasow et al. (2015) have called the "do-be construction", and in the complement of *help*, explored in the present work. Wasow et al. found that a preceding infinitival *to* increases the use of optional following *to* in the environment they examined, but the use of *to* in the complement of *help* is reduced following *to help*. We examine two hypotheses regarding why the same function word is primed by prior use in one construction and inhibited in another. We then test predictions made by the two hypotheses, finding support for one of them.

1 Introduction

Wasow et al. (2015) investigated factors that influence the optional use of *to* in examples like (1),[1] which Flickinger and Wasow (2013) had dubbed the "do-be construction".

All they do is (to) report gloomy things.[2] (1)

The subject of this construction always contains a relative clause containing a form of the verb *do*; its main verb is a copula; and the copula is followed by a verb phrase, whose inflection must take one of three forms: the one matching the form of *do* (as in 2), the full infinitive form (that is, with *to*, as in 3), or a bare infinitive (without *to*, as in 4).

What we're *doing is going* down the same path. (2)
One thing he *did was to listen.* (3)
The best that can be *done is discuss* this issue. (4)

Wasow et al. found that a variety of factors influence the choice between the last two of these. In particular, the rate of *to* in the post-copula VP is significantly higher than would be expected (given the other factors) when the occurrence of *do* in the subject is infinitival, as in (4)—that is, *to do*. This was attributed to the well-known phenomenon of priming (cf. Branigan and Pickering, 2017 and references cited there).

Another environment in which the infinitival *to* is optional is in VP complements of the verb *help*, with or without noun phrase (NP) object, as in (5).

We helped (them) (to) clear the table. (5)

Descriptive grammars of English (e.g., Peters, 2004:247) often note this peculiarity of *help*, sometimes anecdotally suggesting factors that might influence the use of *to*. Among these is the form of *help*. More specifically, in the words of Lohmann's (2011) quantitative corpus study of this phenomenon, "The bare infinitive is preferred after cases of *to help*." This is just the opposite of priming: A preceding *to* reduces, rather than increases, the use of *to* in this construction. Such anti-priming has been given a number of names in the linguistics literature, including *haplology, the Obligatory Contour Principle* (OCP)*,* and *horror aequi.* See Walter (2007) for a detailed discussion and many examples of the application of these terms. We will use the term *interference.*

Our question is why a preceding occurrence of infinitival *to* increases the use of *to* in one environment where it is optional, but has the opposite effect in another. What is it about these two constructions that leads to this difference in the use of *to*?

[1] All examples in this paper are drawn from the Corpus of Contemporary American English (COCA; Davies, 2008-).

[2] The original token included optional *to* in the source corpus.

Proceedings of the Workshop on Cognitive Modeling and Computational Linguistics, pages 11–19
Minneapolis, USA, June 7, 2019.

We begin by presenting a multivariate corpus study of *help (to)*, investigating factors that simultaneously influence the use of *to*, interference being just one among several. Section 2 summarizes the compilation and annotation of our sample then presents our statistical model of these data. Section 3 discusses two possible explanations for the opposite influence of a preceding infinitival *to* on the use of a following optional *to* in the two constructions, then further explores a prediction that follows from one of the two hypotheses presented, providing data confirming that prediction.

2 Corpus Study of *help (to)*

Lohmann's earlier study of *help (to)*, as we have termed the construction, was based on a smaller sample (N=1,718) and explored fewer factors of influence than Wasow et al.'s study of *do-be (to).* The *do-be (to)* work also drew from the Corpus of Contemporary American English (COCA), vs. Lohmann's use of data from the British National Corpus. For better comparison then with the prior results for *do-be (to)*, we have followed Wasow et al. in investigating a similar range of factors, with data drawn from COCA, in the present study using a downloaded version pre-tagged for part of speech, with a total of 520M words divided among five genres: academic, fiction, magazines, newspapers, and spoken.

2.1 Extracting Tokens

We began by programmatically identifying 135K sentences that included a verb-tagged form of *help*. We then passed these sentences through the CoreNLP PCFG constituency parser (version 2.0.2; Klein and Manning, 2003; Manning et al., 2014) to annotate grammatical structure.

Our initial look through several dozen examples found that *help (to)* constructions were represented by a surprising variety of structures in parser output, including a number that we considered to be incorrect. This guided us in crafting a syntactic tree search query (TGrep2; Rohde, 2005) aimed at balancing precision and recall,[3] while still allowing maximal flexibility in terms of any intervening material: between infinitival *to* and *help* (i.e., "split infinitives"); between *help* (or a direct-object NP) and *to* (if present) preceding the complement verb; or between *to* and the complement verb. We identified 78,283 tokens for further analysis. Checking a random sample (N=100) found precision of 98.3%, recall of 76.3%, yielding F_1 measure 85.9.[4]

2.2 Factors in our Analysis

To model variation in our dependent measure, the presence or absence of *to* before a VP complement of *help*, we began by considering elements analogous to those previously shown to be significantly predictive of optional *to* in the *do-be* construction, including phonological, syntactic, cognitive, and information-theoretic measures. Specifically, we programmatically annotated each token for:

- The primary independent variable (predictor) of interest, whether *help* is preceded by infinitival *to* (again, allowing intervening material). Per Lohmann (2011), infinitival *to* is expected to disfavor optional *to* before complement verb (i.e., interference).
- Accessibility of the complement verb lemma, as reflected by relative frequency within the COCA corpus. This was log-adjusted to account for the Zipfian distribution of verb frequencies, as illustrated in Figure 1.

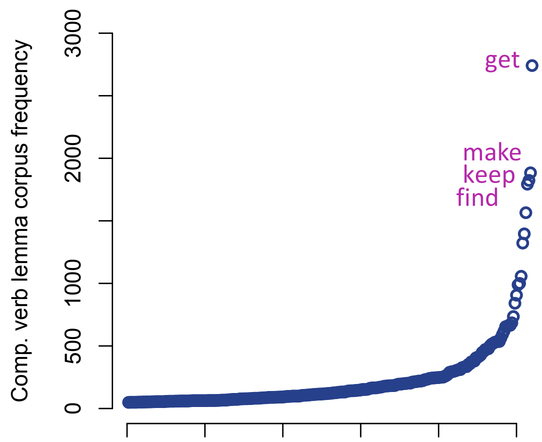

Figure 1: Corpus frequency of *help*-complement verb lemmas. For regression modeling, frequencies are log-adjusted to produce a nearer to linear fit.

[3] Of all tokens returned, *precision* is defined as the fraction that were intended targets. Of all intended targets, *recall* is the portion returned by the search.

[4] We randomly selected 100 tokens from our original sample of all COCA sentences that had included a form of *help* and judged that 76 of these represented *help (to)* constructions. Our subsequent tree search query selected 58 of these valid tokens and one non-*help (to)* token, yielding *precision* = 58/59 (0.983), recall = 58/76 (0.763). F_1 is the harmonic mean of *precision* and *recall*, $2pr/(p+r) = 0.859$.

- Predictability of complement verb lemma in context—context here being its likelihood of following *help (to)*—reflected by its relative frequency within the COCA *help (to)* data, once again log-adjusted.
- Intra-token distances, as derived from the constituency parse, including from head noun of subject NP to complement verb, and from head noun of object NP (if present) to complement verb. Head nouns within NP syntactic constituents were identified via the CoreNLP dependency parser (Chen and Manning, 2014).
- Phonological environment of (optional) *to* site. Where complement verb is preceded by optional *to*, we classified the initial segment of whatever word follows *to* (which may be negation, an intervening adverb, or the complement verb itself) into one of four categories: vowels, sibilants, sonorants, or other.[5] We similarly classify the final segment of whatever word precedes *to*. For tokens omitting optional *to*, we classified the initial segment of the complement verb and the final segment of whatever word precedes it. We then annotated each example for whether the given environment was expected, a priori, to favor or disfavor optional *to*. Since *to* is stop-initial, its insertion was expected to be promoted by OCP when preceded and followed by a pair of vowels, sibilants, or sonorants, but disfavored between pairs of "other" segments (i.e., stops, affricates).
- Stress pattern (i.e., prosody), encoded as *clash*, *lapse*, or *other*. As with phonological environment above, we considered the words following and preceding optional *to* if present, or the complement verb and word preceding it where *to* is omitted. *Clash* was coded if the preceding word has final stress and the following word has initial stress. *Lapse* was coded if these are both unstressed.[6]

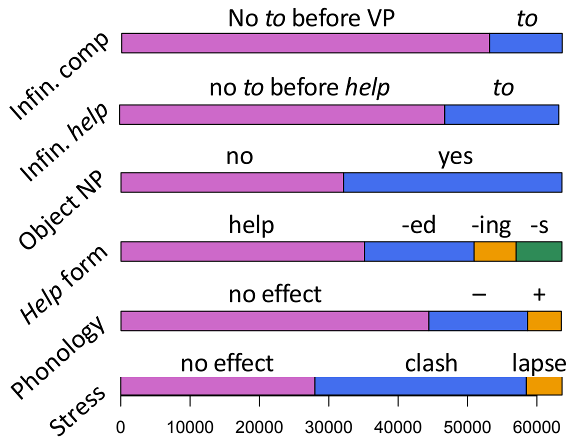

Figure 2: Distribution of categorical variables: (i) presence of infinitival *to* before VP complement; (ii) presence of infinitival *to* before *help*; (iii) presence of direct object NP following *help*; (iv) form of verb *help*; (v) *to*-favoring or disfavoring phonological environment surrounding *to*-site; and (vi) lexical stress environment surrounding *to*-site. X-axis represents number of tokens.

- Surface form of *help* (*help*, *helps*, *helped*, or *helping*).
- Spoken vs. written portion of the corpus.

To these we added measures not modeled in the prior work on the *do-be (to)* construction:

- Animacy of subject.[7]
- Whether or not *help* is negated.
- Whether or not *help* is preceded by a modal auxiliary.

Finally, we encoded an element representing a key difference between *help (to)* and *do-be (to)*:

- Presence of a direct object following *help*.

While the *do-be (to)* construction does not present this option, an object NP following *help* changes the construction's syntactic interpretation. In *help (to)* constructions, *help* is a "control verb", so-called because when followed by a complement verb, *help* functions to control what is understood

[5] Assuming phonetic transcriptions extracted from the Carnegie Mellon Pronouncing Dictionary (CMUdict), version 0.7b (2014), for each word.

[6] Drawing once again on information from CMUdict, in this case syllabic stress for each word.

[7] We follow the programmatic animacy-annotation scheme of Melnick 2017, expanding on a technique from Theijssen 2012. As previously noted, a dependency parse identifies the head noun within each subject NP. This is then lemmatized (via NLTK; Bird et al., 2009) and compared to a static list of animates built from WordNet (Princeton University, 2010) *person* and *animal* terms, a Wikipedia list of notable U.S. companies, and an additional whitelist to capture reflexive pronouns, personal pronouns other than *it* and *them*, and certain impersonal pronouns (*someone*, *everybody*, and so on). Subject head nouns of length greater than two letters in all caps are also marked ANIMATE.

to be the subject of the subordinate VP, but just what that subject is understood to be in any given token depends on whether or not *help* has an object NP. Without an object NP, the subject of the complement verb is understood to be the same as the subject of *help*. For example:

Sunshine helps (to) grow flowers. (6)

In (6), "sunshine" is understood to be the subject of both main verb *help* and complement verb *grow*. In (7), on the other hand, while "sunshine" is again the subject of *help,* object NP "gardeners" is now understood to be the subject of *grow*:

Sunshine helps gardeners (to) grow flowers. (7)

Following annotation, we peformed additional clean-up of the data to improve accuracy. These steps included:

- For the spoken (i.e., transcribed) portion of the corpus only, we excluded tokens where *help* is preceded by *want to*, *have to*, or *going to*, as we suspect that these transcriptions could represent tokens actually spoken closer to a one-word [wɑnə] ("wanna"), [hæftə] ("hafta"), or [gənə] ("gonna"), that is, without a distinct *to* ([tu]), the potential contributor to an interference effect.
- Limiting analysis to tokens with complement verb lemmas appearing 50 or more times within our sample, in order to improve reliability of relative frequency estimates.

After all adjustments, the final data set for analysis totals 63,593 tokens. Figure 2 shows univariate distributions for several factors laid out above.

2.3 Modeling Variation

To assess the effect of infinitival *to* before *help* (i.e., on the inclusion of infinitival *to* before a following complement verb) while simultaneously controlling for other expected influences, we fit our data with a mixed-effects binary logistic regression model (Pinheiro and Bates, 2000; Bresnan et al., 2007; Baayen et al., 2008), predicting infinitival complement VP from fixed effects for the several factors described above, with a random effect for complement verb lemma.

	Est. β	Std Err	z value	Pr(>|z|)	
Fixed effects					
Form of *help*:					
help	-0.608	0.071	-8.614	< 0.0001	***
helped	-0.722	0.082	-8.787	< 0.0001	***
helping	0.418	0.081	5.179	< 0.0001	***
helps	-0.048	0.083	-0.584	0.5590	
Infinitive *help*	**-2.072**	**0.062**	**-33.436**	**< 0.0001**	***
Object NP present ("object control")	-1.691	0.042	-40.741	< 0.0001	***
Written corpus	-0.212	0.034	-6.256	< 0.0001	***
Modal before *help*	0.265	0.044	5.962	< 0.0001	***
Negated *help*	0.413	0.108	3.825	0.0001	***
Subject animacy	-0.300	0.032	-9.417	< 0.0001	***
Phon.: (–) condition	-0.465	0.045	-10.298	< 0.0001	***
Phon.: (+) condition	0.149	0.052	2.881	0.0040	**
Stress: clash	-0.213	0.039	-5.482	< 0.0001	***
Stress: lapse	0.402	0.052	7.707	< 0.0001	***
Distance, controller↔*to*	0.089	0.016	5.488	< 0.0001	***
Verb availaility	0.255	0.044	5.851	< 0.0001	***
Verb predictability	-0.265	0.036	-7.434	< 0.0001	***
Interactions					
Object NP × Sbj animacy	0.218	0.052	4.181	< 0.0001	***
Object NP × Controller distance	0.085	0.022	3.943	0.0001	***
Object NP × Verb preditability	0.250	0.041	6.081	< 0.0001	***

Table 1: Logistic regression model of *help (to)* construction, fixed effects and interactions, predicting optional *to* before complement verb. Positive beta coefficients promote optional *to*. The outlined row highlights the effect of *to* before *help*, with negative coefficient suggesting inhibition (i.e., interference).

Since the presence or absence of an NP direct object following the main verb *help* affects both a given token's projected syntactic structure and its inter-constituent dependencies (e.g., whether the subject of the complement verb is controlled by the subject or object of *help*), we explored interactions of object NP presence with a handful of other predictors, including subject animacy, distance to complement verb from controller (subject or object of *help*), and both availability and predictability of the complement verb. Stepwise reduction based on significant contribution to model fit retained all main effects and eliminated only the interaction of object presence with complement verb availability. Table 1 presents the resulting model, with pseudo-R^2 = 0.382.

The primary observation is that the interference effect of infinitival *help*—i.e., disfavoring optional *to* before a following complement verb—is confirmed here under multivariate control (β = -2.072, z = -33.4, $p < 0.0001$).

While other factors were included chiefly to maximize accuracy of our interference effect estimate, we briefly review their results. Most of the several factors with analogs in Wasow et al.'s model of the *do-be* construction appear to have similar effects here. Written language produces less optional *to* than spoken, presumed to reflect less pressure from online processing demands. Increased distance—in this case, to the complement verb from the subject of *help* or from its direct object, if present—promotes optional *to*, as increased dependency length generates additional processing load (Hawkins, 2004). Increased predictability of a particular complement verb in context (i.e., prior probability of encountering it following *help*) disfavors optional *to*, which we take as an example of the principle of Uniform Information Density (UID; Levy and Jaeger, 2007; Jaeger; 2010). Here, UID would predict that *to* would be more likely to be included where it would serve to spread out the arrival of new information, or *surprisal*, in those cases where the complement verb is *less* predictable in context (i.e., the inverse of predictability). Surprisingly, increased overall corpus frequency of the complement verb—as opposed to its frequency just in the context of the *help (to)* construction—appears to promote optional *to*, counter to its effect in Wasow et al.'s *do-be* results, though exploring this further falls beyond our present scope.

We also find a few significant interactions. The main effect of an animate *help* subject—and thus an animate subject for the complement verb, as well, when *help* has no direct object—appears to disfavor optional *to*, but this effect was largely neutralized in the presence of a direct object. This follows from noting that in such cases, it is the *help* direct object that is interpreted as the subject of the complement verb. The significant distance effect, conversely, was only further enhanced in such "object control" cases. In the context of the shorter dependency length between direct object and following complement verb in these examples, small increases in length had a larger effect. Finally, like the subject animacy effect, the UID effect (predictability of complement verb in context) appears to be largely neutralized in the presence of a direct object.

Figure 3(a) illustrates the relative contribution to model fit for each fixed effect. The presence of an NP object following *help*—with its critical syntactic role, when present, in determining the subject of the complement verb—makes the single largest contribution to model fit, followed by our primary object of study, the interference effect (i.e.,

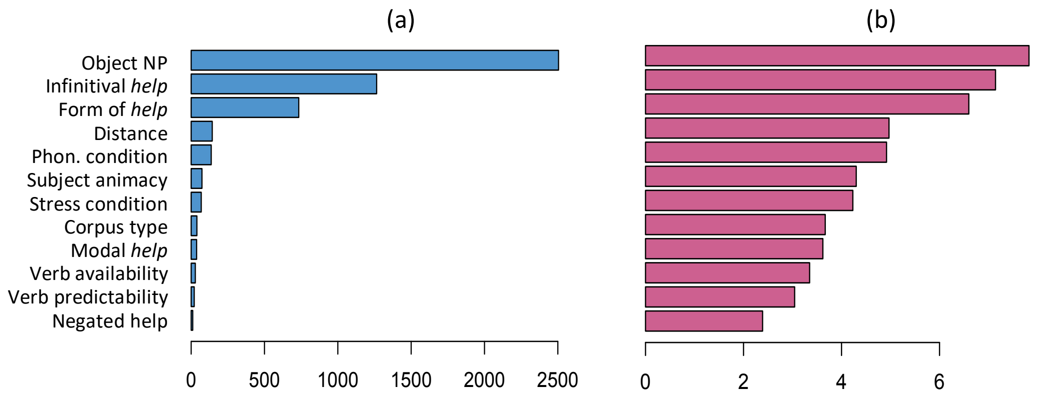

Figure 3: Fixed-effect contributions to fit, as measured by Akaike Information Criterion, log-adjusted in 3(b).

infinitive *to* before *help*). Figure 3(b) presents the same data on a logarithmic scale to better visualize the relative sizes of the smaller contributors.

3 Two Hypotheses

Our model thus confirms under multivariate control the observation that preceding infinitival *to* disfavors optional *to* before a following complement verb, in contrast with the *do-be (to)* construction, where Wasow et al. (2015) had found preceding infinitival *to* favoring optional following *to*—that is, interference in *help (to)* vs. priming in *do-be (to)*. Why do the constructions behave so differently in this respect?

3.1 The Locality Hypothesis

A first hypothesis is that the preceding *to* in the *help* construction tends to be closer to the site of optional *to* than in the *do-be* construction. When no object NP intervenes between infinitival *help* and a VP complement, the site of optional *to* is most often separated from the preceding *to* by just one monosyllabic word. In the *do-be* construction, by contrast, there must be a minimum of two words (*do* and some form of *be*) between infinitival *do* and the site of optional *to*.

Most examples in the linguistics literature of what Walter (2007) calls "repetition avoidance" are very local: avoidance of identical or similar segments, tones, inflections, or words that are adjacent. Hence, it is perhaps natural to conjecture that interference is necessarily a very short-lived effect, and to look for a solution to our puzzle in terms of locality. But the psycholinguistics literature also contains examples of less local interference effects, for example, Ferreira and Firato (2002).

Both our corpus study and that of Wasow et al. found significant effects of the distance to the optional *to* site from an obligatory preceding verb (*do* or *help*). But locality cannot be the full explanation of the difference in the behavior of optional *to* in the two constructions. The interference effect of *to* immediately preceding *help* persists even when an object NP intervenes between *help* and its VP complement. This is confirmed by separately refitting our model to just those tokens with an object NP. The negative influence of preceding *to* on following *to* remains highly significant ($\beta = -1.41$, $z = -17.6$, $p < 0.0001$). In this environment, the optional *to* site following *help* is as far from a preceding *to* as in the *do-be* construction, a minimum of two words in each construction and often more, as in (8). Hence, something else must be involved.

Professionals learn how [*to help*] families of young children with visual impairments [*promote*] emergent literacy skills (8)

3.2 The Function Hypothesis

The second hypothesis was first suggested to us by Emily Bender (p.c.). She noted that the verb *do* that is an obligatory part of the *do-be* construction functions essentially as an elliptical replacement for the post-copula VP. For example, in (1) part of what is predicated of the referent of *they* is that they report gloomy things, and *do* stands in for the VP *report gloomy things*.

Elliptical constructions generally exhibit some structural parallelism between the ellipsis site and the antecedent. The exact nature of the parallelism constraints in such constructions has been the subject of a great deal of linguistic literature over the past half century; see, for example, Hankamer and Sag (1976) and van Craenenbroeck and Merchant (2013). These parallelism constraints presumably assist the listener (or reader) in identifying the antecedent and thus determining the intended interpretation of elliptical expressions. The priming of *to* in the *do-be* construction, then, can be viewed as one component of the expected parallelism in ellipsis.

In contrast, there is no elliptical relationship between the verb *help* and its complement VP. However, when both *help* and its complement are full infinitives (with *to*), it is an instance of self-embedding (also known as recursion)—that is, a construction (in this case, an infinitival VP) directly embedded within another construction of the same type.

It has been known for over half a century (see Miller and Chomsky, 1963:286) that center self-embedding creates severe processing difficulty. Although less attention has been paid to the effect of self-embedding on the edge of a constituent, there is some literature (e.g., Christiansen and MacDonald, 2009) showing that right-branching recursive structures also cause processing difficulty, albeit less than center self-embedding. Without the second occurrence of *to*, *to help* VP is not an instance of self-embedding. Hence, it should not be surprising that we observe interference when *to help* takes a full infinitival complement.

3.3 A Prediction

If the interference effect that we observe in the *help* (*to*) construction is due to avoidance of self-embedding, then it should show up with other verbs that take infinitival complements. Even though *help* is exceptional in allowing the word *to* to be omitted, most other verbs that can take infinitival complements can also occur in other environments. For example, *expect, need, try,* and *want* all can take simple NP objects; *appear, ask,* and *try* can all take a prepositional complement; and *seem* can take an adjectival complement. If speakers avoid embedding infinitival VPs directly under another infinitival VP, the effect should be observable with these other verbs as well. With the other verbs avoiding self-embedding it is not so simple as merely replacing the full infinitival VP with a VP lacking *to*. But other paraphrases that avoid recursion are always possible. Hence, we predict that the rate of occurrence of infinitival VP complements in these other verbs should be lower when the verbs themselves are infinitival (i.e. immediately preceded by *to*) than in other environments.

3.4 Testing the Prediction

To test this prediction, we turn once again to COCA, identifying all verb lemmas ever observed to take an infinitval complement. This yields 10,931 types in the corpus. Further restricting analysis to those verb lemmas appearing more than 1,000 times overall and at least 10 times with an infinitival complement yields 1,019 types. We examined all 70.1M occurances of these verbs, classifying each token into one of four categories: (a) non-infinitival verb, V_1; (b) verb with infinitval complement, V_1 *to* V_2; (c) infinitival verb, *to* V_1; or (d) infinitival verb with infinitval complement, *to* V_1 *to* V_2.[8]

Our prediction can be restated as in (9), the expectation that the conditional probability of the appearance of a complement verb given infinitival main verb should be much less than the conditional probability of complement verb given non-infinitival main verb.

$$p(V_2 \mid to\ V_1) \ll p(V_2 \mid V_1) \qquad (9)$$

This is formulated in (10) in terms of our four-way classification above.

$$\frac{(d)}{(c)} \ll \frac{(b)}{(a)} \qquad (10)$$

In a single metric, we expect the ratio in (11) to be much less than 1.

$$\frac{(d)/(c)}{(b)/(a)} \ll 1 \qquad (11)$$

We found this measure to be less than 1 for 837 of our 1,019 verbs (82.1%), representing 53.4M of 70.1M tokens (76.2%). Across the entire set of verbs, we calculate an aggregate ratio of 0.292. Table 2 presents individual results for ten common verbs that take an infinitival complement, including *help*, in ascending order by ratio value.

Verb	**(d/c) / (b/a)**
help	0.083
hope	0.104
like	0.131
expect	0.133
appear	0.134
seem	0.154
demand	0.232
need	0.386
ask	0.417
try	0.817

Table 2: Values << 1 suggest an infinitival main verb disfavors appearance with an infinitival complement.

To consider the statistical significance of these findings, we employ McNemar's Chi-Squared test, which corrects for lack of independence of observations, required here as each verb provides tokens in multiple conditions (i.e., in each of our four token classification categories). Of the 837 verbs noted above as disfavoring infinitival recursion (ratio < 1), 824 (98.4%) show a significant result under McNemar's $\chi 2$ ($p < 0.05$). Finally, an aggregate test of the full data set finds $\chi 2 = 110.04$, df = 1, $p < 0.0001$.

Infinitival verb self-embedding indeed appears to be strongly disfavored in this large corpus of modern American English, in turn supporting the function hypothesis for the apparent interference effect on optional *to* before a complement verb following infinitival *help*.

[8] Without resorting to parsing the entire 520M-word corpus, we limited extractions here, unlike our full *help* study, to tokens without intervening material.

4 Conclusions

Our corpus investigations of optional *to* have shown that both priming and interference occur in the use of the same optional function word. Which one occurs in a given environment is not arbitrary. Rather, it depends on more general properties of those environments: we find priming where repetition can facilitate processing, as it does in elliptical constructions; and we find interference where repetition creates processing difficulty, as it does in self-embedding.

Acknowledgments

We thank Evan Chuu for assistance in data preparation and are grateful for comments from three anonymous reviewers. This work has also benefited from discussions with Emily Bender, Florian Jaeger, and Maryellen MacDonald. Any errors remain our own. Statistical calculations and graphics were developed with R version 3.5.1 (R Core Team, 2018).

References

R. Harald Baayen, Douglas J. Davidson, and Douglas M. Bates. 2008. Mixed-effects modeling with crossed random effects for subjects and items. *Journal of memory and language* 59.4: 390-412.

Steven Bird, Ewan Klein, and Edwin Loper. 2009. *Natural language processing with Python*. Sebastopol, CA: O'Reilly Media.

Holly P. Branigan and Martin J. Pickering. 2017. Structural priming and the representation of language. *Behavioral and Brain Sciences* 40.

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

Carnegie Mellon University. 2014. *The Carnegie Mellon Pronouncing Dictionary (CMUdict), version 0.7b*. http://www.speech.cs.cmu.edu/cgi-bin/cmudict

Danqi Chen and Christopher Manning. 2014. A fast and accurate dependency parser using neural networks. In *Proceedings of EMNLP 2014*.

Morten H. Christiansen and Maryellen C. MacDonald. 2009. A usage-based approach to recursion in sentence processing. *Language Learning* 59(1): 126-161.

Jeroen van Craenenbroeck and Jason Merchant. 2013. Ellipsis phenomena. In Marcel den Dikken (ed.), *The Cambridge Handbook of Generative Syntax*, 701-745. Cambridge University Press: Cambridge.

Mark Davies. 2008-. *The Corpus of Contemporary American English (COCA): 560 million words, 1990-present*. https://corpus.byu.edu/coca/

Victor S. Ferreira and Carla E. Firato. 2002. Proactive interference effects on sentence production *Psychonomic Bulletin & Review*, 9(4):795-800.

Daniel Flickinger and Thomas Wasow. 2013. A corpus-driven analysis of the do-be construction. In Philip Hofmeister & Elisabeth Norcliffe (eds) *The Core and the Periphery: Data-Driven Perspectives on Syntax Inspired by Ivan A. Sag*, pp. 35-63. CSLI Publications.

Jorge Hankamer and Ivan A. Sag. 1976. Deep and surface anaphora. *Linguistic Inquiry* 7(3):391-426.

John A. Hawkins. 2004. *Efficiency and complexity in grammars*. Oxford University Press on Demand.

Florian Jaeger. 2010. Redundancy and reduction: Speakers manage information density. *Cognitive Psychology* 61(1):23-62.

Dan Klein and Christopher D. Manning. 2003. Accurate unlexicalized parsing. *Proceedings of the 41st Meeting of the Association for Computational Linguistics*, pp. 423-430.

Roger P. Levy and T. Florian. Jaeger. 2007. Speakers optimize information density through syntactic reduction. In *Advances in neural information processing systems*, pp. 849- 856.

Arne Lohmann. 2011. *Help* vs *help to*: A multifactorial, mixed-effects account of infinitive marker omission. *English Language & Linguistics* 15, no. 3: 499-521.

Christopher D. Manning, Mihai Surdeanu, John Bauer, Jenny Finkel, Steven J. Bethard, and David McClosky. 2014. The Stanford CoreNLP Natural Language Processing Toolkit. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, pp. 55-60.

Robin Melnick. 2017. *Consistency in variation: On the provenance of end-weight*. Unpublished Stanford University dissertation.

George Miller and Noam Chomsky.1963. Finitary models of language users. In R. Luce, R. Bush and E. Galanter (eds.) *Handbook of Mathematical Psychology, Vol 2*, 419-93. New York: Wiley.

Pam Peters. 2004. *The Cambridge Guide to English Usage*. Cambridge University Press, Cambridge, England.

José C. Pinheiro and Douglas M. Bates. 2000. *Statistics and computing. Mixed-effects models in S and S-PLUS*. Springer.

Princeton University. 2010. *About WordNet*. http://wordnet.princeton.edu

R Core Team. 2018. *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Douglas L. Rohde. 2005. *Tgrep2 user manual.* http://tedlab.mit.edu/~dr/Tgrep2/tgrep2.pdf

Daphne L. Theijssen. (2012. *Making choices: Modelling the English dative alternation*. Unpublished dissertation, Radboud University Nijmegen.

Mary Ann Walter. 2007. *Repetition avoidance in human language*. Unpublished MIT dissertation.

Thomas Wasow, Roger Levy, Robin Melnick, Hanzhi Zhu, and Tom Juzek. 2015. Processing, prosody, and optional *to*. In Lyn Frazier and Edward Gibson (eds) *Explicit and Implicit Prosody in Sentence Processing*, pp. 133-158. Springer.