

Similarity and Structural Priming

Neal Snider (nsnider@bcs.rochester.edu)

Department of Brain and Cognitive Sciences, Meliora Hall, Box 270268
Rochester, NY 14627-0268 USA

Abstract

The increasing evidence that language processing is sensitive to lexical and structural co-occurrences at different levels of granularity and abstraction (Jurafsky, Bell, Gregory, & Raymond, 2001; Bybee, 2006) has led to the hypothesis that lexical and structural processing may be unified (MacDonald, Pearlmuter, & Seidenberg, 1994; Jurafsky, 1996). This paper examines the specific hypothesis that structural priming and lexical priming may be due to the same underlying mechanism. Lexical priming is known to exhibit sensitivity to the similarity between the prime and the target: the more similar the prime and target words, the greater the magnitude of the priming effect (Ratcliff & McKoon, 1981). Two corpus studies show evidence of an effect of similarity on structural priming. Structural and semantic similarity of the prime and target structures are modeled using a database of ditransitives extracted from the Switchboard corpus and a nearest-neighbor similarity metric. More similar prime and target structures are found to be more likely to occur in the same construction. This effect is in addition to the known similarity effect of verb identity (Pickering & Branigan, 1998), which is controlled through simultaneous multiple regression and model comparison. This suggests that lexical and structural priming could be the same process. Implications for models of representation and processing are discussed.

Keywords: priming; corpus; mixed models; similarity; lexicon; syntax

Introduction

There is increasing evidence that language processing is sensitive to lexical and structural co-occurrences at different levels of granularity and abstraction, and that syntactic knowledge consists of chunks of idiosyncratic linguistic experience that are larger than the word level (Bybee, 2006; Goldberg, 2006; Tomasello, 2003). This is opposed to the traditional “words and rules” view summarized in Pinker (1999), where generalizations about word order and argument structure are stored in a set of rules, and all idiosyncratic information is stored in a lexicon of words. Contrary to the predictions of this view, a series of empirical studies going back over the last couple of decades have shown that people’s behavior is sensitive to the frequency of occurrence of rather large groupings of words, and also that behavior is sensitive to very fine-grained distinctions between stimuli (Jurafsky et al., 2001). These results are also interesting because they indicate that lexical and syntactic processing may be unified in the same model of representation and processing (MacDonald et al., 1994; Jurafsky, 1996).

The purpose of the present study is to investigate the possibility that lexical and syntactic priming are sensitive to the same factors and thus may be the same process operating on the same representational base. The methodology is structural priming. Priming is a particularly useful method for studying the representations involved in lexical and structural

production. Priming is a phenomenon where the processing of a stimulus (the ‘target’) is facilitated if a similar stimulus (the ‘prime’) has been processed previously. This facilitation is greater the more similar the prime and the target, and in fact priming only occurs if they are similar along some cognitive dimension. This is why priming can illuminate the mental representation of knowledge, because if people’s behavior is sensitive to this similarity, then that similarity must arise from the two structures having the same cognitive representation of that dimension. Thus, by experimentally finding the dimensions of similarity between stimuli that cause priming, one can determine the nature of the mental representations of those stimuli. One of the most basic findings in lexical priming is that the magnitude of priming is affected by semantic or associative relatedness. People have less latency in producing a target word following a prime word that is closely related to it (Ratcliff & McKoon, 1981).

This paper presents a pair of corpus studies to investigate whether structural priming of the ditransitive alternation is also sensitive to similarity. The ditransitive alternation involves the choice between the Double Object (DO, 1a) construction, where the recipient argument Noun Phrase (NP) appears before the theme NP, and the Prepositional Object (PO, 1b) construction, where the theme appears before the recipient:

- (1a. *we ... give [a country]_{recipient} [money]_{theme}*¹ (DO)
- b. *we give [money]_{theme} [to a country]_{recipient}*² (PO)

Priming in the ditransitive was demonstrated in the very earliest studies of structural priming (Bock, 1986; Bock & Loebell, 1990) and has been extensively studied in the intervening years (Pickering & Branigan, 1998, etc.). Priming for the ditransitive alternation in conversational speech has also been demonstrated (Gries, 2005; Bresnan, Cueni, Nikitina, & Baayen, 2007).

Like lexical priming, at the structural level, there is some evidence that more similar prime and target structures are more likely to appear in the same construction. Pickering and Branigan (1998) presented a series of production priming experiments using sentence completion. They tested whether various types of structural overlap between prime and target would increase priming. They found that when the same verb is used in both the prime and target, priming is much more likely. They did not find that priming was more likely in their materials when the prime and target verbs shared the same tense, aspect, or agreement. They argued based on these re-

¹Example from the Switchboard corpus

²Hypothetical alternant of (1a)

sults that lexical heads have privileged status in the representation and processing of constructions: construction choice is mediated by choice of the head verb and its argument structure. Another interpretation is that the verb identity boost in priming is due to a more general similarity effect in structural priming, just like that found in lexical priming. The verbal head of a construction is certainly an important part of its representation, especially when it comes to choosing between constructions, because it is a strong cue (Garnsey, Pearlmuter, Myers, & Lotocky, 1997; Bresnan et al., 2007). However, there are other features of each token of a construction that are important for determining construction choice, and these may also affect priming. Other experiments have shown that sharing other features besides the lexical head can increase priming. Cleland and Pickering (2003) showed that in priming adjective-noun ordering in English (*the green square* vs. *the square that's green*), priming is increased when the prime and target share the same head noun, as well as adjective. Further, priming these constructions is increased when the head nouns are semantically similar (in the same semantic class), although not when they are phonologically similar. Also Griffin and Weinstein-Tull (2003) showed that Verb Phrase complement constructions are more likely to be repeated when they are similar with respect to having the same number of semantic arguments. All these results are consistent with the hypothesis that structural priming is sensitive to a general similarity between constructions. The purpose of the current paper is to test whether similarity in structural priming can be measured using a continuous metric of general similarity between prime and target structures, just like lexical priming.

The Data section below describes the data used for the priming studies: a database of ditransitives drawn from a corpus of naturalistic dialogue. Study 1 replicates in a corpus study the Pickering and Branigan (1998) experimental finding that priming is boosted when the verb is repeated. Study 2 presents a metric for measuring structural-semantic similarity between prime and target structures, and tests whether this affects priming beyond the similarity due to verb repetition. Finally, the paper concludes with a discussion of the implications for models of representation and processing.

Data

The data that will be used to examine similarity effects in structural priming comes from a corpus study of the ditransitive alternation. Bresnan et al. (2007) presented an extensive study of the ditransitive alternation and the factors that influence the choice of PO or DO construction. They extracted all the alternating verbs in the Treebank Switchboard (Marcus, Santorini, & Marcinkiewicz, 1994) that occurred in either the DO or PO construction. They then searched through the entire Switchboard corpus, not just the part parsed in the Penn Treebank project, and extracted the alternating verbs that they had identified as ditransitive in the Treebank. They also annotated their database with many features of the construction that they hypothesized might affect choice of ar-

gument ordering, for example the semantic class of the verb and the givenness and animacy of the recipient and theme. They used a logistic regression analysis to determine which of these factors were significant predictors of the ditransitive alternation. Bresnan et al. found that nonpronominal, inanimate, nongiven, and indefinite recipients are more likely to be in the PO construction. On the other hand, nongiven and indefinite themes are more likely to be in the DO construction. Also, if the theme was longer than the recipient, a DO is more likely. Finally, a previous use of PO in the dialogue made PO more likely, so priming is evident by this measure. This same database from Bresnan et al. (2007) was used for the current study of the factors that affect of priming in the ditransitive, and the factors mentioned above were used as controls. The probability that the verb occurs in the PO (verb bias) was also added because Bresnan et al. (2007) found that some verbs are biased towards one alternant. These controls are very important when studying structural priming in dialogue, because there are many other reasons that constructions might be repeated (Szmrecsányi, 2005), such as stylistic choice, where two tokens are in the same construction because they are in similar information-structural environments. In an attempt to isolate the persistence effects that are due to priming, the factors from the Bresnan et al. (2007) will be controlled when the effects of interest are reported.

The first step in this study of priming in the ditransitive is to add a factor that indicates whether the previous prime construction was PO or DO (as opposed to the more general parallelism factor in Bresnan et al. (2007), which only indicated if a DO had occurred previously in the dialogue). In this study and those that follow, only priming-related factors are of interest, so the subset of the data will be used where a prime construction is measurable (the first token in a dialogue will have no prime). This yields a database of 1002 tokens, of which 215 are PO (the whole Bresnan et al. database has 2349 tokens, of which 499 are PO).

Study 1: Verb repetition and structural priming

If structural priming behaves like lexical priming, then one would predict that structural priming is sensitive to similarity between the prime and target, just as in lexical priming. The results of the experiments by Pickering and Branigan (1998) indicate that prime constructions are more likely to be repeated when the prime and target share the same verb. In order to test the prediction that this result is due to a broader similarity effect on structural priming using the ditransitive database, it must first be demonstrated that priming in such natural dialogue is more likely when the prime and target tokens have the same verb.

Methods

The head verb of the prime construction was extracted for each token in the database, and a factor was added that indicated whether the prime and target tokens had the same verb. To determine whether the interaction between prime

construction and prime-target verb identity is statistically significant, a logistic regression model similar to that in Bresnan et al. (2007) was used, with likelihood of PO being the positive response. Specifically, a mixed logit model³ was used in both studies in order to control for a random effect of speaker (in other words, each speaker is allowed to have a different rate of ditransitive choice). Mixed logit models can be thought of as an extension of logistic regression that includes modeling of random effects. Inclusion of random effects is necessary to generalize beyond the speakers in the data sample. In this study, the random effect accounts for the fact that speakers may have different rates of ditransitive use. Here and in all following studies, the coefficient β for each independent variable and its levels of significance will be reported. Coefficients in logistic regression models are given in log-odds (the space in which logistic models are fitted to the data). The dependent variable was structure choice, DO or PO.

Statistical tests on the model are also reported. The Wald's Z tests whether the log-odds coefficient of the factor is significantly different from zero, and therefore tests the directionality of the coefficient, whether it is positive, and biases towards the PO, or negative, and biases towards the DO. In order to determine whether the factor is significantly predictive in the model (regardless of direction), a likelihood ratio test ($\Delta_x(\Lambda)$) is performed. To perform this test, a model is run with all the factors included in order to determine the log-likelihood of the data according to the full model. Then, to test for the significance of a factor x , that factor is removed and it is noted how much the data log-likelihood changed according to the model. The $\Delta_x(\Lambda)$ test is also more robust against collinearity in the model, so even though the factors are highly correlated (like the main effect of priming and its interaction with the factor of interest in this study, verb identity), one can determine which is actually driving the effects on ditransitive choice.

Results

Figure 1 illustrates graphically the number of persistent constructions with respect to verb repetition. The height of the boxes on the y-axis indicates the proportion of prime target pairs for which the same construction was used in each condition, and the width of the boxes on the x-axis indicates the proportion of the data that occurred in each condition. The table shows that a greater proportion of prime construction repetitions occurs when the verb is the same between prime and target, as opposed to when the verbs are different.

The interaction between priming and verb identity was added to the model (which includes all the factors from Bresnan et al. (2007) as controls), and was found to significantly increase priming ($\beta = 1.1489$, odds = 3.15, $SE = 0.6494$, $Z = 1.769$, $p_Z < .08$, $\chi^2 = 3.2$, $p_{\chi^2} < .08$)⁴. A PO prime is 3.15 times more likely to be repeated if the prime and target share the same verb.

³the *lme4* library in the *R* statistical package

⁴This result also appeared as a control factor in Jaeger and Snider (2008)

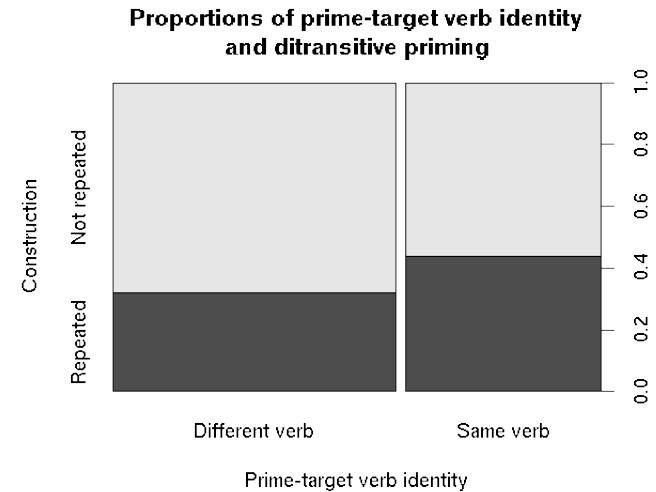


Figure 1: Spine plot illustrating the effect of prime-target verb identity on ditransitive choice.

Discussion

The results of this study replicate the experimental findings of Pickering and Branigan (1998): priming is increased when the prime and target share the same verb. Pickering and Branigan interpreted this result as evidence for the privileged status of lexical heads in the representation and processing of constructions: construction choice is mediated by choice of the head verb and its argument structure. Another interpretation is that the verb identity boost in priming is due to a more general similarity effect in structural priming, just like that found in lexical priming. The verbal head of a construction is certainly an important part of its representation, especially when it comes to choosing between constructions, because it is a strong cue (Garnsey et al., 1997; Bresnan et al., 2007). However, there are other features that are important for determining construction choice, and these may also affect similarity and priming. The next study investigates this possibility.

Study 2: Structural priming and similarity

If lexical and structural processing are similar because they draw on similar representations, then one would expect to find that the processing mechanisms involved in structural priming will be affected by the same factors as lexical priming, like structural and semantic similarity. This study tests the hypothesis that structural priming is increased when the prime and target are more similar, that is, when they share more features than just their head verb.

Methods

The main methodological issue then for this study is how to measure the similarity between prime and target tokens. Nearest Neighbor (NN) exemplar models (Daelemans, Zavrel, Sloot, & Bosch, 2001) provide just such a similarity

measure. NN models classify cases by comparing the current case with all other observed exemplars (usually a data set derived from a corpus), and picking the best classification based on past experience. An essential part of this model is the distance metric used to determine how similar exemplars are to one another. The algorithm uses this to weight highly the exemplars that are closest to the case to be classified. However, such a distance metric could also be extremely useful in determining how similar one structure is to another for other purposes, such as predicting their priming behavior.

Exemplars in NN models are defined as sets of features. The distance between exemplar X and exemplar Y , $\Delta(X, Y)$, is the weighted sum of the difference between the two exemplars along each of the dimensions defined by the features:

$$\Delta(X, Y) = \sum_i w_i \delta(x_i, y_i) \quad (1)$$

where w_i is the weight associated with feature i , and $\delta(x_i, y_i)$ is defined as:

$$\delta(x_i, y_i) = \begin{cases} \frac{|x_i - y_i|}{\max_i - \min_i} & \text{if numeric, otherwise} \\ 0 & \text{if } x_i = y_i \\ 1 & \text{if } x_i \neq y_i \end{cases}$$

If the feature is categorical, the the difference between the two exemplars is 0 if they have the same value for the feature, and 1 if they have different values for the feature. If the feature is continuous, then the difference between the two exemplars is just the difference in the values of feature i for the two exemplars, divided by the maximum range of that feature (so that features won't be weighted too highly just because they involve larger numbers).

Applying such a model to priming, the hypothesis for this study is that two exemplars that are more similar in the sense that they share more features and have a lower NN-distance between them, are more likely to prime. The similarity hypothesis predicts that the interaction between the prime construction and the NN-distance will be negative: the less distance between the prime and target structures, the more likely the target will be in the same construction as the prime. To test the hypothesis, the NN-distance according to equation (1) was calculated between every prime-target pair in the ditransitive database. The features used to calculate the NN-distance are all the features that were included in the Bresnan et al. (2007) study, that is, all the non-priming-related factors in Table 1, plus features representing the head verb, the head of the recipient NP, and the head of the theme NP. The feature weights were absolute values of the coefficients in a control model that was run which did not include priming factors. These coefficients/weights did not differ significantly from the values in the second column of Table 1. As an example, two tokens that have a small exemplar distance (0.66) are the prime *He hasn't given it to someone else* and target *I'll take it to him*, which, although they have different verbs, both have inanimate, pronominal themes and animate, pronominal recipients. On the other hand, the prime *I'm there*

to teach them something and target *I don't pay much attention to it* have a larger distance of 1.2, which reflects the fact that they only share a pronominal recipient, but differ on all other dimensions (different verbs, the prime theme is pronominal while the target theme is abstract and non-pronominal, etc.).

Results

To demonstrate graphically the trends in the data with respect to prime-target similarity, Figure 2 shows how the NN-distance relates to proportions of POs produced when the prime construction is PO (note this graph only reflects trends in data, and does not reflect the effect of this factor while controlling for other possibly confounding factors). The height of

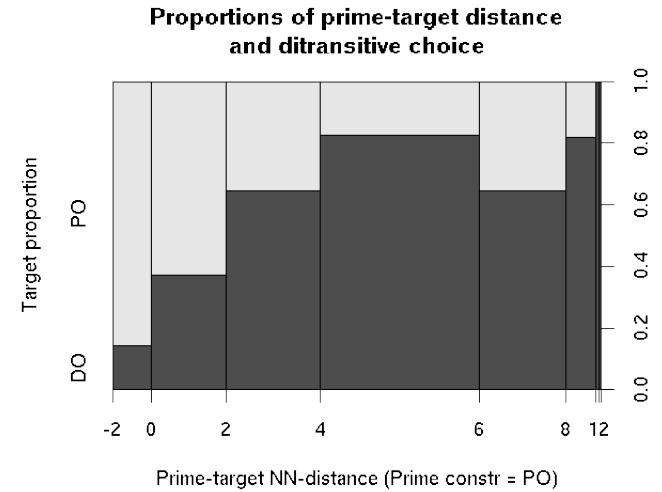


Figure 2: Spine plot illustrating the the effect of prime-target similarity on ditransitive choice.

the boxes on the y-axis show the proportion of PO and DO targets for each similarity bin when the prime construction was PO. The width of the bins on the x-axis shows how much of the data falls in each bin of the similarity factor. The graph arguably shows a negative trend: more POs are produced when the construction is PO and the k -NN distance between prime and target is low than when the prime-target distance is high.

To determine whether the trend in figure Figure 2 is statistically significant, a factor was added to the model representing the interaction between the **Prime construction** and **NN-distance**. Also, the factor from Study 1 was added to control for the the boost of priming when the prime and target verb are the same (Pickering & Branigan, 1998). Thus the model will test whether the similarity metric based on all features of the prime and target structures is a better predictor of priming than a predictor based on verb identity alone. This tests the prediction of the similarity hypothesis that more similar primes and targets should prime more, and the similarity is not merely head-driven: similarity is a continuous effect based on the associative strength between two structures. Thus, the crucial test for the similarity effect on priming is

Table 1: Summary of statistical analysis of the effect of prime-target similarity on ditransitive choice

Predictor (independent variable)	Parameter estimates			Wald's test		$\Delta_x(\Lambda)$ -test	
	Log-odds	S.E.	Odds	Z	P	χ^2	P
Nonpronominal recipient	0.34	0.46	1.8	1.3	$\ll 0.001$	1.7	< 0.2
Inanimate recipient	3.57	0.51	36	6.9	< 0.2	48	$\ll 0.001$
Nongiven recipient	2.32	0.50	10	4.7	$\ll 0.001$	22	$\ll 0.001$
Indefinite recipient	0.78	0.51	2.1	1.5	< 0.2	2.2	< 0.15
Plural theme	0.93	0.36	2.9	3.0	< 0.005	8.7	< 0.005
Nongiven theme	-1.25	0.43	0.28	-3.0	< 0.005	8.9	< 0.005
log(length difference)	-1.37	0.24	0.24	-6.0	$\ll 0.001$	35	$\ll 0.001$
Indefinite theme	-2.26	0.41	0.09	-6.0	$\ll 0.001$	35	$\ll 0.001$
Nonpronominal theme	-1.6	0.42	0.18	-4.0	$\ll 0.001$	16	$\ll 0.001$
Semantic class	-2.8	0.50	0.07	-5.4	$\ll 0.001$	29	$\ll 0.001$
Verb bias	0.53	0.10	1.7	5.1	$\ll 0.001$	26	$\ll 0.001$
Prime constr = PO	0.65	0.32	1.9	2.1	< 0.05	4.3	< 0.05
Prime constr x Identity	-0.024	0.82	0.97	-0.1	< 1	0.03	< 1
Prime constr x NN-dist	-2.36	0.87	0.05	-2.7	< 0.01	10.8	< 0.005

that, while prime and target exemplars that share a verb are certainly more similar, having identical verbs does not drive the whole effect. The model comparisons ($\Delta_x(\Lambda)$) will test whether the effect of similarity still influences priming even when prime-target verb identity is controlled.

The results of the statistical analysis are presented in Table 1. The coefficients in log-odds and standard errors associated with the control factors are given in the second and third column. The corresponding odds coefficients are given the fourth column. The fifth and sixth columns summarize the Wald's Z statistic, which tests whether the coefficients are significantly different from zero. Finally, the last two columns give the χ^2 over the change in data likelihood ($\Delta_x(\Lambda)$) associated with the removal of the predictor (x) from the final model. There is a clearly significant interaction between prime construction and the NN-distance between the prime and target exemplars such that the less distance between the prime and target, the more likely they are to both be in the same construction. When the prime construction is PO, the PO construction is 10.6 times more likely in the target for every one-unit decrease in the distance between the prime and target in feature space. This effect is in addition to the known verb repetition boost in structural priming because the NN-distance explains all the variation when both it and the verb identity factor are in the model. However, verb identity is still a strong factor in determining similarity because **both** verb identity ($p < .005$) and NN-distance ($p < .05$) interact with priming when the verb identity feature is excluded from the NN-distance metric.⁵

Discussion

This study provides evidence that prime-target similarity significantly strengthens structural priming: target tokens that share more features with the prime are more likely to re-

peat the prime construction. Further, this effect is not merely driven by the head of the construction (the verb), although this is certainly a strong contributor to similarity.

These results indicate that structural priming is affected by a general notion of structural-semantic similarity, like lexical priming. Thus, lexical and structural production are affected by priming in the same way with respect to similarity, which is consistent with a model of representation and production in which there is no explicit distinction between lexical and structural storage.

Conclusion

These studies present evidence from modeling ditransitive choice in natural language corpora that structural priming is more likely the more features the prime and target structures share. This has a clear implication for the study of structural priming, that the verb repetition boost in structural priming found by Pickering and Branigan (1998) is a specific case of a more general similarity effect in priming. This similarity hypothesis also makes the further prediction that overlap of other words and features should boost priming in single-factor priming experiments like the Pickering and Branigan and Cleland and Branigan studies. Experimentally testing this hypothesis would also strengthen the causal link between similarity and priming, to establish that it is indeed similarity that causes persistence and not choice of the same structure that causes similar words to be chosen.

More generally, these studies provide further evidence that lexical and syntactic priming are affected by similar factors. They complement results presented by Jaeger and Snider (2008), who presented evidence that structural priming is affected by the predictability of the prime such that the prime construction is more likely to be repeated the less predictable it is given the verb. A similar “inverse predictability effect” was found in lexical priming in that people require less time

⁵More details about this study may be found in Snider (2008)

to react to target words (producing them and deciding if they are words) if the target was preceded by a low frequency prime than if it was preceded by a high frequency prime (Scarborough, Cortese, & Scarborough, 1977; Perea & Rosa, 2000).

One implication of these results is that models of processing, particularly production, may need to explain why lexical and structural production mechanisms are affected by similar factors. Production models usually have separate storage mechanisms for lexical information and structural information (Bock, 1999), but these models could incorporate the representations from Usage-based grammars (Bybee, 2006; Goldberg, 2006; Tomasello, 2003), which make available linguistic representations with different degrees of abstraction and granularity, and do not have an explicit representational distinction between word and structure. An example of a representational model that uses such structures, which have the abstraction necessary for the priming results presented here, but still have fine lexical detail necessary to capture the similarity effects is the Data-Oriented Parsing model of Bod (1992). This model uses all possible substructures of a corpus to parse new strings, so it implicitly discovers constructions of various grain sizes, including single words. A production model that uses such representations would have no explicit distinction between words and structures, and could capture these results.

References

Bock, J. (1986). Syntactic persistence in language production. *Cognitive Psychology*, 18(3), 355–387.

Bock, J. (1999). Language production. In R. Wilson & F. Keil (Eds.), *Mit encyclopedia of the cognitive sciences* (pp. 453–456). Cambridge, MA: MIT Press.

Bock, J., & Loebell, H. (1990). Framing sentences. *Cognition*, 35(1), 1–39.

Bod, R. (1992). Data Oriented Parsing (DOP). *Proceedings COLING'92*.

Bresnan, J., Cueni, A., Nikitina, T., & Baayen, H. (2007). Predicting the dative alternation. *Proceedings of the Royal Netherlands Academy of Science Workshop on Foundations of Interpretation*. Amsterdam.

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

Cleland, A. A., & Pickering, M. J. (2003). The use of lexical and syntactic information in language production: Evidence from the priming of noun-phrase structure. *Journal of Memory and Language*, 49, 214–230.

Daelemans, W., Zavrel, J., Sloot, K. van der, & Bosch, A. van den. (2001). *TiMBL: Tilburg memory based learner, version 4.0, reference guide* (Tech. Rep.). ILK Technical Report 01-04, Tilburg University. available from <http://ilk.kub.nl>.

Garnsey, S. M., Pearlmuter, N. J., Myers, E., & Lotocky, M. A. (1997). The contributions of verb bias and plausibility to the comprehension of temporarily ambiguous sentences. *Journal of Memory and Language*, 37, 58–93.

Goldberg, A. (2006). *Constructions at Work: The Nature of Generalization in Language*. Oxford University Press.

Gries, S. (2005). Syntactic Priming: A Corpus-based Approach. *Journal of Psycholinguistic Research*, 34(4), 365–399.

Griffin, Z., & Weinstein-Tull, J. (2003). Conceptual structure modulates structural priming in the production of complex sentences. *Journal of Memory and Language*, 49(4), 537–555.

Jaeger, T. F., & Snider, N. (2008). Implicit learning and syntactic persistence: Surprise and cumulativity. In *Proceedings of the 30th annual conference of the cognitive science society*.

Jurafsky, D. (1996). A probabilistic model of lexical and syntactic access and disambiguation. *Cognitive Science*, 20, 137–194.

Jurafsky, D., Bell, A., Gregory, M., & Raymond, W. D. (2001). Probabilistic relations between words: Evidence from reduction in lexical production. In J. L. Bybee & P. J. Hopper (Eds.), *Frequency and the emergence of linguistic structure* (pp. 229–254). Amsterdam: John Benjamins.

MacDonald, M. C., Pearlmuter, N., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review*, 101, 676–703.

Marcus, M. P., Santorini, B., & Marcinkiewicz, M. A. (1994). Building a large annotated corpus of English: The Penn Treebank. *Computational Linguistics*, 19(2), 313–330.

Perea, M., & Rosa, E. (2000). Repetition and form priming interact with neighborhood density at a brief stimulus onset asynchrony. *Psychonomic Bulletin and Review*, 7(4), 668–77.

Pickering, M., & Branigan, H. (1998). The Representation of Verbs: Evidence from Syntactic Priming in Language Production. *Journal of Memory and Language*, 39(4), 633–651.

Pinker, S. (1999). *Words and Rules: The Ingredients of Language*. Basic Books.

Ratcliff, R., & McKoon, G. (1981). Does activation really spread. *Psychological Review*, 88(5), 454–462.

Scarborough, D., Cortese, C., & Scarborough, H. S. (1977). Frequency and repetition effects in lexical memory. *Journal of Experimental Psychology: Human Perception and Performance*, 3, 1–17.

Snider, N. (2008). *An exemplar model of syntactic priming*. Unpublished doctoral dissertation, Stanford University, Stanford, CA.

Szmrecsányi, B. M. (2005). Language users as creatures of habit: a corpus-based analysis of persistence in spoken English. *Corpus Linguistics and Linguistic Theory*, 1, 113–149.

Tomasello, M. (2003). *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Harvard University Press.