Measuring Syntactic Priming in Dialog Corpora

Armin Buch and Christian Pietsch
Universität Tübingen and Universität Bielefeld
armin.buch@uni-tuebingen.de,cpietsch+le@uni-bielefeld.de

Abstract

We devise a simple, distribution-based measure of priming between linguistic categories. Priming is found in tree banks of dialog corpora, both for context-free production rules and for Combinatory Categorial Grammar categories. It is stronger for task-oriented dialogs, and stronger in lexical categories than in syntactic categories.

1 Introduction

Priming[1] is the phenomenon by which a recently encountered event (or seen item) is recognized more quickly or more likely to be repeated. Presumably this is due to activation levels in the brain. Production and perception prime each other, as they access the same or closely related representations. *Direct* priming only acts on surface similarity. The according effect for learned relations is called *associative* or *semantic* priming. Only the latter is affected by age, amnesia etc., so it is an effect of memory (Tulving and Schacter, 1990). One expects it to decay exponentially over time.

In language, hearing or speaking a word facilitates the processing of similar sounding words. The so-called semantic priming arises between semantically similar concepts, between syntactic categories etc. Priming has been attested for in single constructions such as the English dative alternation (Bock and Griffin, 2000). Any identifiable unit in a linguistic structure could be subject to priming. If it is, this supports the linguistic theory which proposed that structure. “[R]epeatable structures are evidence for the units of linguistic cognition” (Reitter, 2008, sec. 1.2). Priming of syntactic categories has been found by Reitter et al. (2006a,b).

[1] This research has been supported by the EC 277 “Cognitive Interaction Technology” of the DFG at Bielefeld University. We would like to thank David Reitter and Julia Hockenmaier for kindly providing us with their corpora, and our adviser Gerhard Jäger for creating and guiding this project.

2 Experiments

2.1 Preliminaries

Classical priming experiments such as Bock and Griffin (2000) study a single, theory-neutral alternation in controlled experiments. In contrast, we study the distribution of each category in large annotated syntactic corpora. Every sub-structure of an annotation is a possible category. We follow Reitter et al. (2006a,b) in using Combinatory Categorial Grammar (CCG) categories and context-free production rules. The latter could be extended to subtrees as in data-oriented parsing (Bod, 1998).

CCG assumes that there are many equivalent derivations for a given sentence analysis: the same lexical categories, but different modes of combination. Among these, the *normal form* derivation is the one along the lines of constituent bracketing, which is mostly right-branching for languages such as English. The *incremental* derivation is as left-branching as possible; see Reitter et al. (2006a) for details.

We use the same data as Reitter et al. (2006a,b): The Switchboard corpus, annotated with context-free rule expansions (**sw-CFG**) and with CCG categories (**sw-CCG-I** and **sw-CCG-N** for incremental and normal form derivations, respectively; and the MapTask corpus with CFG annotation (**mp-CFG**). We also look at lexical priming (**sw-words**).

2.2 A simple measure of priming

Priming as mental activation of representations cannot be directly measured (yet). In corpus studies, we observe the distribution of a category. The null hypothesis is a random distribution, described as a Poisson process. For this, the (temporal) distances between adjacent occurrences are exponentially distributed ($p(x) = \lambda_0 e^{-\lambda_0 x}$), where λ_0 equals the frequency of the category.

We fit an exponential curve with decay parameter λ to the actual distribution of distances. If there is priming, shorter distances should be more frequent than longer distances ($\lambda > \lambda_0$). The ratio $r = \lambda / \lambda_0$ can be interpreted as priming strength.

2.3 Single Categories

Fig. 1 shows the estimated density function, a random distribution (dotted line), and the fitted, much steeper exponential (dashed line), for the expansion **VP → VB S**.

Across all corpora, estimated parameters λ are always larger than λ_0. Rare categories show more priming with r up to 2.3, and close to 1 for the very common expansion S → NP VP (0.34 occurrences per second). Exponential decay fits well, with standard deviation around 0.005.

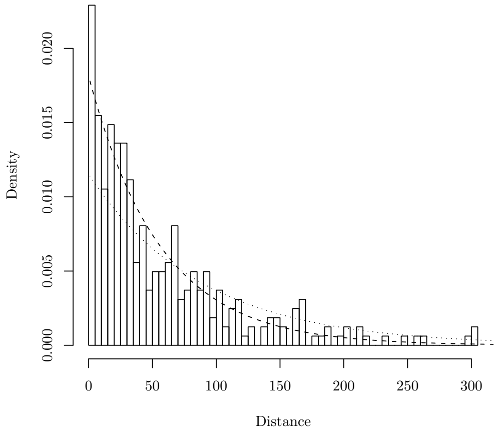

Figure 1: Distribution of pairwise distances of VP → VB S.

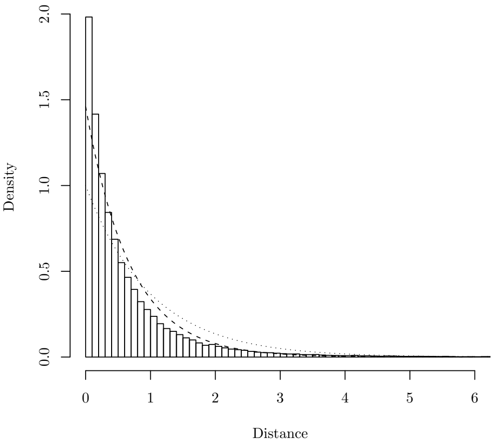

Figure 2: Average of normalized distributions in MapTask.

The exponential decay supports the suggestion that priming is an effect of (short-term) memory. While frequent categories have generally less room for skewed distributions, there is still something more to be explained about the effect of frequency. Besides that, our results once more confirm the existence and pervasiveness of priming.

2.4 Corpus averages

Measuring the overall priming in a corpus allows to compare several settings: different linguistic frameworks (CCG vs. CFG), spoken vs. written language, conversational (Switchboard) vs. task-oriented (MapTask). We normalize all categories for frequency (s.t. $\lambda_0 = 1$) and take the average.

Corpus	decay parameter λ	standard error
sw-CFG	1.1589	0.0044
sw-CCG-I	1.0523	0.0054
sw-CCG-N	1.0364	0.0051
mp-CFG	1.4666	0.0049
sw-words	1.2521	0.0113

Standard errors are low, we have thus a good estimate of the actual distribution of distances. Yet fig. 2 suggests an even more extreme distribution. This might be a result of cumulating activation: short distances trigger more short distances.

We see strong lexical priming (1.25). Task-oriented dialog outranks conversational dialog (Reitter et al., 2006b; Pickering and Garrod, 2004). CCG annotation shows comparably little priming. Results by Reitter et al. (2006a) stated that it is significant, but that the difference is not.

3 Conclusion

We have devised a notably simple priming measure. A single parameter λ per category (or per corpus) suffices, modeling the distribution of distances. Experiments show it to be larger than its expected value, which is the category's frequency. The effect appears to be larger for rare categories. Interpreting the fitted λ as a frequency is somewhat paradoxical: Primed categories appear more frequent than they actually are.

So far we have viewed categories as mutually exclusive. This does not take into account priming of *similar* categories. Adding pairwise similarities to the model could improve it. A simple example is stemming or lemmatization: A word also primes all inflected forms.

The more priming a category shows, the more it can be taken as psycholinguistically valid. We plan to use priming to evaluate grammar formalisms and their proposed categories. As pointed out by Reitter (et al.), this is a novel approach to inform linguistic theory (about linguistic *competence*) by *performance* data.

References

Bock, K. and Z. M. Griffin (2000). The persistence of structural priming: Transient activation or implicit learning? *Journal of Experimental Psychology*, **129**(2):177–192.

Bod, R. (1998). *Beyond Grammar: An Experience-Based Theory of Language*. CSLI Lecture Notes. CSLI Publications, Stanford. ISBN 157586150X.

Pickering, M. J. and S. Garrod (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, **27**(02):169–190.

Reitter, D. (2008). *Context Effects in Language Production: Models of Syntactic Priming in Dialogue Corpora*. Ph.D. thesis, University of Edinburgh.

Reitter, D., J. Hockenmaier, and F. Keller (2006a). Priming effects in Combinatory Categorial Grammar. In *Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 308–316. Sydney, Australia.

Reitter, D., J. D. Moore, and F. Keller (2006b). Priming of syntactic rules in task-oriented dialogue and spontaneous conversation. In *Proceedings of the 28th Annual Conference of the Cognitive Science Society (CogSci)*, pp. 685–690. Vancouver, Canada.

Tulving, E. and D. L. Schacter (1990). Priming and human memory systems. *Science*, **247**(4940):301–306.