Computational Modelling of Structural Priming in Dialogue

David Reitter, Frank Keller, Johanna D. Moore
dreitter | keller | jmoore @ inf.ed.ac.uk
School of Informatics
University of Edinburgh
United Kingdom

Abstract

Syntactic priming effects, modelled as increase in repetition probability shortly after a use of a syntactic rule, have the potential to improve language processing components. We model priming of syntactic rules in annotated corpora of spoken dialogue, extending previous work that was confined to selected constructions. We find that speakers are more receptive to priming from their interlocutor in task-oriented dialogue than in sponaneous conversation. Low-frequency rules are more likely to show priming.

1 Introduction

Current dialogue systems overlook an interesting fact of language-based communication. Speakers tend to repeat their linguistic decisions rather than making them from scratch, creating *entrainment* over time. Repetition is evident not just on the obvious lexical level: *syntactic* choices depend on preceding ones in a way that can be modelled and, ultimately, be leveraged in parsing and language generation. The statistical analysis in this paper aims to make headway towards such a model.

Recently, priming phenomena[1] have been exploited to aid automated processing, for instance in automatic speech recognition using cache models, but only recently have attempts been made at using them in parsing (Charniak and Johnson, 2005). In natural language generation, repetition can be used to increase the alignment of human and computers. A surface-level approach is possible by biasing the n-gram language model used to select the output string from a variety of possible utterances (Brockmann et al., 2005).

Priming effects are common and well known. For instance, speakers access lexical items more quickly after a semantically or phonologically similar prime. Recent work demonstrates large effects for particular synonymous alternations (e.g., active vs. passive voice) using traditional laboratory experiments with human subjects (Bock, 1986; Branigan et al., 2000). In this study, we look at the effect from a computational perspective, that is, we assume some form of parsing and syntax-driven generation components. While previous studies singled out syntactic phenomena, we assume a phrase-structure grammar where all syntactic rules may receive priming. We use large-scale corpora, which reflect the realities of natural interaction, where limited control exists over syntax and the semantics of the utterances. Thus, we quantify priming for the general case in the realistic setting provided by corpus based experiments. As a first hypothesis, we predict that after a a syntactic rule occurs, it is more likely to be repeated shortly than a long time afterwards.

From a theoretical perspective, priming opens a peephole into the architecture of the human language faculty. By identifying units in which priming occurs, we can pinpoint the structures used in processing. Also, priming may help explain the ease with which humans engange in conversations.

This study is interested in the differences relevant to systems implementing language-based human-

[1] The term *priming* refers to a process that influences linguistic decision-making. An instance of priming occurs when a syntactic structure or lexical item giving evidence of a linguistic choice (*prime*) influences the recipient to make the same decision, i.e. re-use the structure, at a later choice-point (*target*).

Proceedings of the Human Language Technology Conference of the North American Chapter of the ACL, pages 121–124,
New York, June 2006.

computer interaction. Often, HCI is a means for user and system to jointly plan or carry out a task. Thus, we look at repetition effects in task-oriented dialogue. A recent psychological perspective models *Interactive Alignment* between speakers (Pickering and Garrod, 2004), where mutual understanding about task and situation depends on lower-level priming effects. Under the model, we expect priming effects to be stronger when a task requires high-level alignment of situation models.

2 Method

2.1 Dialogue types

We examined two corpora. *Switchboard* contains 80,000 utterances of *spontaneous spoken conversations* over the telephone among randomly paired, North American speakers, syntactically annotated with phrase-structure grammar (Marcus et al., 1994). *The HCRC Map Task* corpus comprises more than 110 dialogues with a total of $20,400$ utterances (Anderson et al., 1991). Like Switchboard, HCRC Map Task is a corpus of spoken, two-person dialogue in English. However, Map Task contains *task-oriented dialogue*: interlocutors work together to achieve a task as quickly and efficiently as possible. Subjects were asked to give each other directions with the help of a map. The interlocutors are in the same room, but have separate, slightly different maps and are unable to see each other's maps.

2.2 Syntactic repetitions

Both corpora are annotated with phrase structure trees. Each tree was converted into the set of phrase structure productions that license it. This allows us to identify the repeated use of rules. Structural priming would predict that a rule *(target)* occurs more often shortly after a potential *prime* of the same rule than long afterwards – any repetition at great distance is seen as coincidental. Therefore, we can correlate the probability of repetition with the elapsed time (DIST) between prime and target.

We considered very pair of two equal syntactic rules up to a predefined maximal distance to be a potential case of priming-enhanced production. If we consider priming at distances $1 \ldots n$, each rule instance produces up to n data points. Our binary response variable indicates whether there is a prime for the target between $n - 0.5$ and $n + 0.5$ seconds before the target. As a prime, we see the invocation of the same rule. Syntactic repetitions resulting from lexical repetition and repetitions of unary rules are excluded. We looked for repetitions within windows (DIST) of $n = 15$ seconds (Section 3.1).

Without priming, one would expect that there is a constant probability of syntactic repetition, no matter the distance between prime and target. The analysis tries to reject this null hypothesis and show a correlation of the effect size with the type of corpus used. We expect to see the syntactic priming effect found experimentally should translate to more cases for shorter repetition distances, since priming effects usually decay rapidly (Branigan et al., 1999).

The target utterance is included as a random factor in our model, grouping all 15 measurements of all rules of an utterance as *repeated measurements*, since they depend on the same target rule occurrence or at least on other other rules in the utterance, and are, thus, partially inter-dependent.

We distinguish *production-production priming* within (PP) and *comprehension-production priming* between speakers (CP), encoded in the factor ROLE. Models were estimated on joint data sets derived from both corpora, with a factor SOURCE included to discriminate the two dialogue types.

Additionally, we build a model estimating the effect of the raw frequency of a particular syntactic rule on the priming effect (FREQ). This is of particular interest for priming in applications, where a statistical model will, all other things equal, prefer the more frequent linguistic choice; recall for competing low-frequency rules will be low.

2.3 Generalized Linear Mixed Effect Regression

In this study, we built generalized linear mixed effects regression models (GLMM). In all cases, a rule instance *target* is counted as a repetition at distance d iff there is an utterance *prime* which contains the same rule, and *prime* and *target* are d units apart. GLMMs with a logit-link function are a form of *logistic regression*.[2]

[2]We trained our models using Penalized Quasi-Likelihood (Venables and Ripley, 2002). We will not generally give classical R^2 figures, as this metric is not appropriate to such GLMMs. The below experiments were conducted on a sample of 250,000

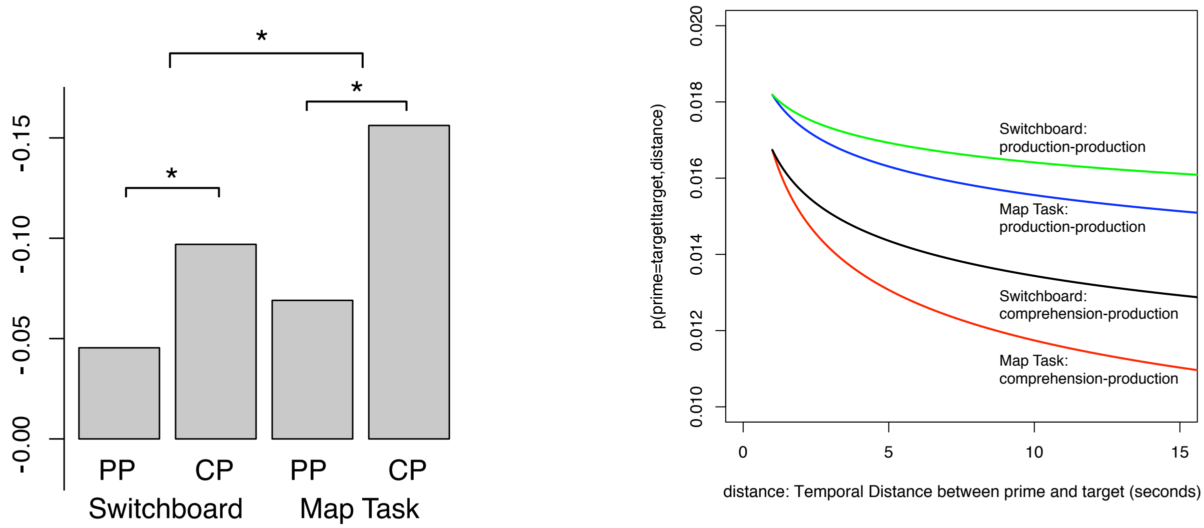

Figure 1: Left: Estimated priming strength (repetition probability decay rate) for Switchboard and Map Task, for within-speaker (PP) and between-speaker (CP) priming. Right: Fitted model for the development of repetition probability (y axis) over time (x axis, in seconds). Here, decay (slope) is the relevant factor for priming strength, as shown on the left. These are derived from models without FREQ.

Regression allows us not only to show that priming exists, but it allows us to predict the decline of repetition probability with increasing distance between prime and target and depending on other variables. If we see priming as a form of pre-activation of syntactic nodes, it indicates the decay rate of preactivation. Our method quantifies priming and correlates the effect with secondary factors.

3 Results

3.1 Task-oriented and spontaneous dialogue

Structural repetition between speakers occured in both corpora and its probability decreases logarithmically with the distance between prime and target.

Figure 1 provides the model for the influence of the four factorial combinations of ROLE and SOURCE on priming (left) and the development of repetition probability at increasing distance (right). SOURCE=Map Task has an interaction effect on the priming decay $ln(\text{DIST})$, both for PP priming ($\beta = -0.024, t = -2.0, p < 0.05$) and for CP priming ($\beta = -0.059, t = -4.0, p < 0.0005$). (Lower coefficients indicate more decay, hence more priming.)

data points per corpus.

In both corpora, we find positive priming effects. However, PP priming is stronger, and CP priming is much stronger in Map Task.

The choice of corpus exhibits a marked interaction with priming effect. Spontaneous conversation shows significantly less priming than task-oriented dialogue. We believe this is not a side-effect of varying grammar size or a different syntactic entropy in the two types of dialogue, since we examine the *decay of repetition probability* with increasing distance (interactions with DIST), and not the overall probability of chance repetition (intercepts / main effects except DIST).

3.2 Frequency effects

An additional model was built which included $ln(\text{FREQ})$ as a predictor that may interact with the effect coefficient for $ln(\text{DIST})$.

$ln(\text{FREQ})$ is inversely correlated with the priming effect (Paraphrase: $\beta_{lnDist} = -1.05, \beta_{lnDist:lnFreq} = 0.54$, Map Task: $\beta_{lnDist} = -2.18, \beta_{lnDist:lnFreq} = 0.35$, all $p < 0.001$). Priming weakens with higher (logarithmic) frequency of a syntactic rule.

4 Discussion

Evidence from Wizard-of-Oz experiments (with systems simulated by human operators) have shown that users of dialogue systems strongly align their syntax with that of a (simulated) computer (Branigan et al., 2003). Such an effect can be leveraged in an application, provided there is a priming model interfacing syntactic processing.

We found evidence of priming in general, that is, when we assume priming of each phrase structure rule. The priming effects decay quickly and non-linearly, which means that a dialogue system would best only take a relatively short preceding context into account, e.g., the previous few utterances.

An important consideration in the context of dialogue systems is whether user and system collaborate on solving a task, such as booking tickets or retrieving information. Here, syntactic priming *between* human speakers is strong, so a system should implement it. In other situations, systems do not have to use a unified syntactic architecture for parsing and generation, but bias their output on previous system utterances, and possibly improve parsing by looking at previously recognized inputs.

The fact that priming is more pronounced *within* (PP) a speaker suggests that optimizing parsing and generation separately is the most promising avenue in either type of dialogue system.

One explanation for this lies in a reduced cognitive load of spontaneous, everyday conversation. Consequently, the more accessible, highly-frequent rules prime less.

In task-oriented dialogue, speakers need to produce a common situation model. Interactive Alignment Model argues that this process is aided by syntactic priming. In support of this model, we find more priming in task-oriented dialogue.[3]

5 Conclusions

Syntactic priming effects are reliably present in dialogue even in computational models where the full range of syntactic rules is considered instead of selected constructions with known strong priming.

This is good news for dialogue systems, which tend to be task-oriented. Linguistically motivated systems can possibly exploit the user's tendency to repeat syntactic structures by anticipating repetition. Future systems may also align their output with their recognition capabilities and actively align with the user to signal understanding. Parsers and realizers in natural language generation modules may make the most of priming if they respect important factors that influence priming effects, such as task-orientation of the dialogue and frequency of the syntactic rule.

[3]For a more detailed analysis from the perspective of interactive alignment, see Reitter et al. (2006).

Acknowledgements

The authors would like to thank Amit Dubey, Roger Levy and Martin Pickering. The first author's work is supported by a grant from the Edinburgh Stanford Link.

References

A. Anderson, M. Bader, E. Bard, E. Boyle, G. M. Doherty, S. Garrod, S. Isard, J. Kowtko, J. McAllister, J. Miller, C. Sotillo, H. Thompson, and R. Weinert. 1991. The HCRC Map Task corpus. *Language and Speech*, 34(4):351–366.

J. Kathryn Bock. 1986. Syntactic persistence in language production. *Cognitive Psychology*, 18:355–387.

Holly P. Branigan, Martin J. Pickering, and Alexandra A. Cleland. 1999. Syntactic priming in language production: Evidence for rapid decay. *Psychonomic Bulletin and Review*, 6(4):635–640.

Holly P. Branigan, Martin J. Pickering, and Alexandra A. Cleland. 2000. Syntactic co-ordination in dialogue. *Cognition*, 75:B13–25.

Holly P. Branigan, Martin J. Pickering, Jamie Pearson, Janet F. McLean, and Clifford Nass. 2003. Syntactic alignment between computers and people: the role of belief about mental states. In *Proceedings of the Twenty-fifth Annual Conference of the Cognitive Science Society*.

Carsten Brockmann, Amy Isard, Jon Oberlander, and Michael White. 2005. Modelling alignment for affective dialogue. In *Workshop on Adapting the Interaction Style to Affective Factors at the 10th International Conference on User Modeling (UM-05)*. Edinburgh, UK.

Eugene Charniak and Mark Johnson. 2005. Coarse-to-fine n-best parsing and MaxEnt discriminative reranking. In *Proc. 43th ACL*.

M. Marcus, G. Kim, M. Marcinkiewicz, R. MacIntyre, A. Bies, M. Ferguson, K. Katz, and B. Schasberger. 1994. The Penn treebank: Annotating predicate argument structure. In *Proc. ARPA Human Language Technology Workshop*.

Martin J. Pickering and Simon Garrod. 2004. Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, 27:169–225.

David Reitter, Johanna D. Moore, and Frank Keller. 2006. Priming of syntactic rules in task-oriented dialogue and spontaneous conversation. In *Proceedings of the 28th Annual Conference of the Cognitive Science Society*.

William N. Venables and Brian D. Ripley. 2002. *Modern Applied Statistics with S. Fourth Edition.* Springer.